AP, see amalgamation property
EQ, see elimination of quantifiers
JEP, see joint embedding property
T[A], see Definition 3.10
Diag, see diagram
E(·), see existentially closed structure
Md(·), see model
S(·), see submodel
Th(·), see theory
=, see isomorphism
≡, see elementarily equivalent
≡₀, see elementarily equivalent
≡₁, see elementarily equivalent
≡₂, see elementarily equivalent
<, see elementary substructure
<₀, see elementary substructure
<₁, see elementary substructure
∃-type, 121
maximal, 122
of a point in a structure, 121
∃₁ max(T,v), see Definition 6.32
∀ₙ-axiomatizable, see Definition 3.14
≡([∃ₙ], see Definition 3.2
≡([∀ₙ], see Definition 3.2
|=, see satisfaction
|=, see semantic consequence
|=₀, see proof theoretic consequence
0-tame
sentence, 112
amalgamation property, 63
assignment
for a structure, 13
atomic
structure for a theory, 138
theory, 141
axioms
for a theory, 18
b&f, see back-and-forth
back-and-forth system, 73
batch of formulas, 12
cardinality
of a language, 7
of a structure, 4
carrier
of a structure, 4
chain
of structures, 53
compactness theorem
crude version, 23
refined version, 25
companion
model, 108
operator, 111
least example, 114
theories, 107
complete
theory, 19
consequence
proof theoretic, 20
semantic, 19
consistent
set of sentences, 23
crude compactness theorem, 23
dense linear order without end points, 31
diagram
of a structure, 45
directed
family of structures, 53
downwards Löwenheim-Skolem, 58
elementarily equivalent, 16, 41
elementary
class of structures, 18
elementary substructure, 42
Eli Bers, 111
elimination of quantifiers, 29
embedding
  between structures, 42
equivalent
  formulas relative to a theory, 29
existentially closed structure, 114
existentially universal structure, 126

finitely satisfiable
  set of sentences, 23
formula
  complete over a theory, 137
full type, 121
  of a point in a structure, 121
generic structure
  for infinite forcing, 131
graph
  a kind of structure, 83
    random, 84

infinite forcing, 131
isomorphic structures, 10
isomorphism, 10
  partial, 72
joint embedding property, 63

Löwenheim-Skolem theorems, 58
line
  dense linear order without end points, 31
linear order types, 77
  arithmetic of, 77
    product of, 77
    sum of, 77
mates
  nodes of a graph, 83
maximal-∃-type, 122
model
  of a set of sentences, 18
    of a theory, 48
model companion, 108
model complete theory, 61
model completion, 111
modelling relation, 14
omitting a type, 122

partial isomorphism, 72
point
  of a structure, 16
potentially isomorphic, 73
product of linear order types, 77
proof theoretic consequence, 20
quantifier complexity, 10
quantifier elimination, 29
quantifier hierarchy, 11
random graph, 84
realizing a type, 122
refined compactness theorem, 25
satisfaction relation, 14
satisfiable
  set of sentences, 23
saturated structure, 125
semantic consequence, 19
signature, 4
strictly elementary
  class of structures, 18
structure, 4
  ∃₁-atomic for a theory, 138
    atomic, 138
    existentially closed, 114
    existentially universal, 126
generic
  for infinite forcing, 131
saturated, 125
substructure, 9
submodel
  of a theory, 48
submodel complete theory, 68
substructure, 9
  elementary, 42
sum of linear order types, 77
theory
  ∃₁-atomic, 141
    atomic, 141
categorical, 19
complete, 19
model complete, 61
of a class of structures, 18
submodel complete, 68
two successor functions
theory of, 135

type
  $\exists_1$, 121
  a set of formulas, 12, 121
  full, 121
  of a point in a structure, 121
  omitting a, 122
  over a structure, 125
  realizing a, 122

upwards Löwenheim-Skolem, 58
Basic Model Theory

Harold Simmons

A first course in model theory

This is a first draft and will be updated as the course proceeds.
This version was compiled 2 March 2008
# Contents

## 1 Development

1 Syntax and semantics
   1.1 Signature and language .................................. 4
       Exercises ............................................. 8
   1.2 Basic notions ........................................... 9
       Exercises ............................................. 12
   1.3 Satisfaction ............................................. 13
       Exercises ............................................. 17
   1.4 Consequence ............................................. 18
       Exercises ............................................. 21
   1.5 Compactness ............................................. 23
       Exercises ............................................. 26

2 The effective elimination of quantifiers .................. 29
   2.1 The generalities of quantifier elimination ............... 29
       Exercises ............................................. 31
   2.2 Linear orders ........................................... 31
       Exercises ............................................. 34
   2.3 The natural numbers ..................................... 34
       Exercises ............................................. 39
   2.4 Some other examples – To be done ......................... 40
       Exercises ............................................. 40

3 Basic methods ............................................. 41
   3.1 Some semantic relations .................................. 41
       Exercises ............................................. 43
   3.2 The diagram technique ................................... 44
       Exercises ............................................. 49
   3.3 Restricted axiomatization ................................ 49
       Exercises ............................................. 52
   3.4 Directed families of structures .......................... 53
       Exercises ............................................. 57
   3.5 The up and down techniques .............................. 58
       Exercises ............................................. 60
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Start Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Model complete and submodel complete theories</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>Model complete theories</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>63</td>
</tr>
<tr>
<td>4.2</td>
<td>Two structural properties</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>Submodel complete theories</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>The back and forth technique</td>
<td>71</td>
</tr>
<tr>
<td>5.1</td>
<td>The technique</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>76</td>
</tr>
<tr>
<td>5.2</td>
<td>Linear orders</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>83</td>
</tr>
<tr>
<td>5.3</td>
<td>Graphs</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>90</td>
</tr>
<tr>
<td>5.4</td>
<td>Equivalence relations</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>94</td>
</tr>
<tr>
<td>5.5</td>
<td>Miniature arithmetic</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>106</td>
</tr>
<tr>
<td>5.6</td>
<td>Perhaps algebraically closed abelian groups if I can remember the details</td>
<td>106</td>
</tr>
<tr>
<td>5.7</td>
<td>Survey – of other examples</td>
<td>106</td>
</tr>
<tr>
<td>6</td>
<td>Companion theories and existentially closed structures</td>
<td>107</td>
</tr>
<tr>
<td>6.1</td>
<td>Model companions</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>111</td>
</tr>
<tr>
<td>6.2</td>
<td>Companion operators</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>114</td>
</tr>
<tr>
<td>6.3</td>
<td>Existentially closed structures</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>117</td>
</tr>
<tr>
<td>6.4</td>
<td>Existence of existentially closed structures</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>120</td>
</tr>
<tr>
<td>6.5</td>
<td>The use of types</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>123</td>
</tr>
<tr>
<td>7</td>
<td>Large and small structures</td>
<td>125</td>
</tr>
<tr>
<td>7.1</td>
<td>Existentially universal structures</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>129</td>
</tr>
<tr>
<td>7.2</td>
<td>A companion operator</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>132</td>
</tr>
<tr>
<td>7.3</td>
<td>Existence of existentially universal structures</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>136</td>
</tr>
<tr>
<td>7.4</td>
<td>Atomicity</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>144</td>
</tr>
</tbody>
</table>
# Contents

## II Solutions

A Syntax and semantics
- A.1 Signature and language ............................................. 147
- A.2 Basic notions ....................................................... 153
- A.3 Satisfaction ......................................................... 154
- A.4 Consequence ......................................................... 158
- A.5 Compactness ......................................................... 163

B The effective elimination of quantifiers
- B.1 The generalities of quantifier elimination ..................... 167
- B.2 Linear orders ....................................................... 167
- B.3 The natural numbers ................................................ 169
- B.4 Some other examples—**not yet done** ............................. 173

C Basic methods
- C.1 Some semantic relations ........................................... 175
- C.2 The diagram technique ............................................. 175
- C.3 Restricted axiomatization ......................................... 177
- C.4 Directed families of structures ................................... 179
- C.5 The up and down techniques ..................................... 180

D Model complete and submodel complete theories
- D.1 Model complete theories .......................................... 181
- D.2 Two structural properties ......................................... 184
- D.3 Submodel complete theories ...................................... 186

E The back and forth technique
- E.1 The technique ....................................................... 187
- E.2 Linear orders ....................................................... 189
- E.3 Graphs .............................................................. 190
- E.4 Equivalence relations .............................................. 193
- E.5 Miniature arithmetic .............................................. 193

F Companion theories an existentially closed stuctures
- F.1 Model companions ................................................ 195
- F.2 Companion operators ............................................. 196
- F.3 Existentially closed structures ................................... 196
- F.4 Existence of existentially closed structures .................. 198
- F.5 The use of types .................................................. 200

G Large and small structures
- G.1 Large and small structures ...................................... 203
- G.2 Large and small structures ...................................... 203
- G.3 For §7.3—**to be done** ......................................... 204
- G.4 Atomicity ......................................................... 204
Part I
Development
Model theory (and, in fact, much of mathematical logic) is concerned, in part, with the use of syntactic gadgetry (languages) to extract information about the objects under investigation. To do this efficiently we must have (at least at the back of our mind) a precise definition of the underlying language and the associated facilities. Setting up such a definition can look complicated, and is certainly tedious. However, the ideas involved are essentially trivial. The role of the definition is merely to delimit what can, and what can’t, be done in a first order or elementary language.

In the next section we look at the details of this definition, and in the rest of this chapter we look at some of the consequential ramifications. Before we do that let’s motivate the ideas behind the constructions.

A structure $\mathcal{A}$ (of the kind we deal with in model theory) is a non-empty set $A$, called the carrier of the structure, furnished with some distinguished attributes each of which is an element of $A$, a relation on $A$, or an operation on $A$. Many objects used in mathematics are structures in this sense, but many are not. The crucial restriction here is that the structure is single sorted with just one carrier, this carrier is non-empty, and all the furnishings are first order over the carrier. In the wider scheme of things these are serious restrictions.

It is a curiosity that in model theory a structure is always non-empty. However, for certain situations, the study of relational structures for instance, it would make more sense to allow empty structures. The root of this eccentricity is historical. The standard accounts of predicate calculus don’t handle empty structures very well, and at one time model theory was somehow intertwined with that. We now know better.

Often we are concerned with a whole family of structures each of the same similarity type or signature. For such a family there is a common language which can be used to talk about any and all of these structures. This language is generated in a uniform way from the signature.

This language $L$ allows the use of the connectives

\[ \text{not} \quad \text{and} \quad \text{or} \quad \text{implies} \]

(and perhaps others of the same ilk). It also allows some quantification. However, in any instance of a quantification the bound variable can only, and must, range over the carrier (of the structure being talked about). The language allows quantification over elements of the carrier, but not over subsets of, lists from, or any other kind of gadget extracted from the carrier. This is the first order restriction.

In more advanced work other kinds of languages are used, but not here. At this level model theory is about the use of first order languages, and the exploitation of a central and distinguishing result, the compactness theorem.
1. Syntax and semantics

1.1 Signature and language

In this section we make precise the ideas outlined in the preamble. This will take several steps but, as explained, there is nothing very complicated going on.

1.1 DEFINITION. A signature is an indexed family of symbols where each is either

- a constant symbol $K$
- a relation symbol $R$ with a nominated arity
- an operation symbol $O$ with a nominated arity

respectively. Each arity is a non-zero natural number. We speak of an $n$-placed relation symbol or an $n$-placed operation symbol to indicate the the arity is $n$.

The word ‘symbol’ here indicates that eventually we will generate a formal language $L$ of certain finite strings. The letters ‘$K$’, ‘$R$’, and ‘$O$’ have been chosen in a rather awkward manner to remind us that they are syntactic symbols.

Before we start to generate the full language $L$ let’s take a quick look at the kind of gadget that will provide the semantics for $L$.

1.2 DEFINITION. A structure $\mathfrak{A}$ for a given signature consists of the following.

- A non-empty set $A$, the carrier of $\mathfrak{A}$.
- For each constant symbol $K$, a nominated element $\mathfrak{A}[K]$ of $A$.
- For each $n$-placed relation symbol $R$, a nominated $n$-placed relation $\mathfrak{A}[R]$ on $A$.
- For each $n$-placed operation symbol $O$, a nominated $n$-placed operation $\mathfrak{A}[O]$ on $A$.

These nominated gadgets are the distinguished attributes of $\mathfrak{A}$.

You should differentiate between the symbols $K$, $R$, and $O$ of the signature and the interpretation $\mathfrak{A}[K]$, $\mathfrak{A}[R]$, and $\mathfrak{A}[O]$ of these in the particular structure $\mathfrak{A}$. There is only one language $L$ of the given signature, but there are many different structures of that signature, each of which provides an interpretation of each symbol of the signature.

There are times in model theory when we need to take note of the size of a structure (or a language).

1.3 DEFINITION. The cardinality $|\mathfrak{A}|$ of a structure is the cardinality $|A|$ of its carrier.

Notice that a signature can be empty, in which case a structure (for that signature) is just a non-empty set. At the other extreme, a signature can be uncountably large. On the whole, in this book we are concerned with signatures that have no more than countably many relation symbols and operation symbols. However, for technical reasons, it is convenient to allow the number of constant symbols to be arbitrarily large.

1.4 DEFINITION. Given a signature, the associated primitive symbols are as follows.
1.1. Signature and language

- The symbols of the signature.
- The equality symbol \( \approx \).
- An unlimited stock of variables \( v \).
- The connectives \( \neg, \land, \lor, \rightarrow \).
- The quantifiers \( \forall \) and \( \exists \).
- The constant sentences which are true and false.
- The punctuation symbols ( and ).

There are no other primitive symbols.

A **string** is a finite list of primitive symbols.

In other words, the primitive symbols of the language \( L \) are the symbols of the signature together with a fixed collection of other symbols. These extra symbols are the same for each language.

Again, you should not confuse the formal symbol ‘\( \approx \)’ (which is a primitive of each language) with the informal symbol ‘\( = \)’ used to indicate the equality of two gadgets.

A string is any finite list of primitive symbols, and can be complete gibberish

\[
))\neg v \approx (\land w) \rightarrow \exists
\]

or can be potentially meaningful

\[
(\forall v)(\exists w)((fvw \approx v) \land \neg(fwv \approx w))
\]

where \( v, w \) are variables and \( f \) is a 2-placed operation symbol. Our aim is to extract the potentially meaningful strings.

We do that in three steps. We define the terms \( t \), the atomic formulas \( \theta \), and then the formulas \( \phi \). The formulas are the potentially meaningful strings. Each of these strings

\[
t \quad \theta \quad \phi
\]

has an associated support

\[
\partial t \quad \partial \theta \quad \partial \phi
\]

giving the set of variables occurring freely in the string. The support is generated at the same time as the parent string.

1.5 **Definition.** Each signature has an associated set of terms and each such term \( t \) has an associated set \( \partial t \) of free variables. These are generated as follows.

- Each variable \( v \) is a term, and \( \partial v = \{ v \} \).
- Each constant symbol \( K \) is a term, and \( \partial K = \emptyset \).
• For each \( n \)-placed operation symbol \( O \) and each list \( t_1, \ldots, t_n \) of terms, the compound

\[(Ot_1 \cdots t_n)\]

is a term and

\[\partial(Ot_1, \cdots t_n) = \partial t_1 \cup \cdots \cup \partial t_n\]

is its set of free variables.

There are no other terms.

1.6 **DEFINITION.** For a signature the atomic formulas are those strings

\[
\text{true} \quad \text{false} \quad (t_1 \equiv t_2) \quad (Rt_1 \cdots t_n)
\]

where \( t_1, t_2, \ldots t_n \) are terms and \( R \) is an \( n \)-placed relation symbol. Each such atomic formula \( \theta \) has a set \( \partial \theta \) of free variables given by

\[
\emptyset \quad \emptyset \quad \partial t_1 \cup \partial t_2 \quad \partial t_1 \cup \cdots \cup \partial t_n
\]

respectively.

Notice the difference between

\[(t_1 = t_2) \quad (t_1 \equiv t_2)\]

where \( t_1, t_2 \) are terms. The first asserts that the two terms are the same (that is, the same string of primitive symbols), whereas the second is an atomic formula of the language which, in isolation, has no truth value.

Finally we can extract the potentially meaningful strings.

1.7 **DEFINITION.** Each signature has an associated set of formulas and each such formula \( \phi \) has an associated set \( \partial \phi \) of free variables. These are generated as follows.

- Each atomic formula is a formula with free variables, as given.
- For each formula \( \psi \) the string \( \neg \psi \) is a formula with \( \partial \neg \psi = \partial \psi \).
- For each pair \( \theta, \psi \) of formulas, each of the strings

\[
(\theta \land \psi) \quad (\theta \lor \psi) \quad (\theta \rightarrow \psi)
\]

is a formula with free variables

\[\partial \theta \cup \partial \psi\]

in each case.
For each formula $\psi$ and variable $v$, each of the strings

$$(\forall v)\psi \quad (\exists v)\psi$$

is a formula with free variables

$$\partial \psi - \{v\}$$

in both cases.

There are no other formulas.

Notice that once again each formula can be uniquely parsed.

More importantly, notice how

$$\partial (\forall v)\psi \quad \partial (\exists v)\psi$$

are calculated. The variable $v$ may be free in $\psi$, that is $v \in \partial \psi$, but it becomes bound in

$$(\forall v)\psi \quad (\exists v)\psi$$

and is no longer free. This means that a variable can be used more than once in a formula in essentially different roles. Have a look at Exercise 1.1.

The language $L$ given by a signature is the set of all formulas associated with that signature. In practice we usually don’t even mention the signature, but use phrases such as

structure suitable for the language $L$

term of the language $L$

and so on.

In a way the formulas are not the most important strings associated with a signature.

1.8 DEFINITION. A sentence of a language is a formula $\sigma$ of that language with no free variables, $\partial \sigma = \emptyset$.

Sentences are those strings which are either true or false in any particular structure. Formulas are really just a stepping stone in the construction of sentences.

It may not be clear what role the two constant sentences true and false play.

At times it is convenient to have a sentence which is trivially valid in every structure we meet and, as a string, is very simple. If the language has a constant $K$, then $(K \equiv K)$ is such a sentence. However, sometime there isn’t a constant around, and then we have to look elsewhere. We could take the sentence $(\forall v)(v \equiv v)$, but that quantifier can be a nuisance. The primitive true is there so that we always have such a trivially true and simple sentence. The primitive false plays a similar role, except this one is a trivially false and simple sentence.

Finally, for this section, as indicated above we want to measure the size of a language.

1.9 DEFINITION. The cardinality $|L|$ of a language $L$ is either $\aleph_0$ or the size of the signature, whichever is the larger.
The cardinal $|L|$ is infinite. If the signature is finite or countable, then $|L| = \aleph_0$, otherwise it is the cardinality of the signature. For the most part we will be interested in countable languages. However, even for such a language, one of the most common techniques of model theory involves the use of associated languages of larger cardinality. Thus we have to deal with the general case.

In the following exercises you will see various references to the compactness theorem. This is discussed in Section 1.5. You should come back to these exercises once you are familiar with that notion.

**Exercises**

1.1 Consider the signature with just one symbol $<$ which is a 2-placed relation symbol. We write this as an infix. Let $u, v$ be a pair of distinct variables and set

$$\phi_0 := (v \equiv v) \quad \theta_r := (\exists v)[(u \equiv v) \land \phi_r] \quad \phi_{r+1} := (\exists u)[(u < v) \land \theta_r]$$

for each $r < \omega$ to obtain two $\omega$-chains of formulas.

(a) Write down $\phi_0, \phi_1, \phi_2,$ and perhaps $\phi_3$ until you can see what is going on.

(b) Calculate $\partial \phi_r$ and $\partial \theta_r$ for each $r < \omega$.

(c) Describe a different way of setting up ‘equivalent’ formulas which makes the uses of the variables easier to see.

1.2 Consider the signature with just one relation symbol, which is 2-placed, and no other symbols.

(a) Write down axioms which characterize the classes of reflexive relations, irreflexive relations, symmetric relations, antisymmetric relations, transitive relations, equivalence relations, pre-orders, partial orders.

Observe the quantifier complexity of each sentence you use.

Do any of these notions need infinitely many axioms to characterize them?

(b) Write down axioms which characterize those equivalence relations such that for each strictly positive natural number $n$ there is precisely one block (equivalence class) of size $n$.

How many sentences do you need?

Can you find extra sentences which characterize those structures in this class for which there are no infinite blocks.

1.3 Write down a signature that determines a language which is suitable for describing properties of monoids, commutative monoids, groups, abelian groups, divisible abelian groups.

In your language write down sets of sentences which characterize each of these classes. Observe the quantifier complexity of the sentences you use.

How do you handle inverses in groups?

What about the classes of groups in which each non-neutral element has infinite order, finite order, prime order, respectively?

1.4 Write down a signature that determines a language which is suitable for describing properties of rings of various kinds.
1.2. Basic notions

In your language write down sets of sentences which characterize all rings, all integral domains, all fields, all fields of characteristic $p$ (for a given prime $p$), all fields of characteristic zero, algebraically closed fields (of a given characteristic).

What about the class of fields of non-zero characteristic.

Observe the shape and quantifier complexity of the sentences you use. Can you spot something?

1.5 Consider the natural numbers $\mathbb{N}$ with some appropriate furnishings. Recall that the Peano axioms (discovered by Dedekind) characterize the structure up to isomorphism. Using an appropriate first order language investigate the possible characterizing axioms.

1.6 Consider rationals $\mathbb{Q}$ and the reals $\mathbb{R}$. We can view each of these in three different ways; as a linearly ordered set, as a field, as a linearly ordered field.

For each of these views try to find sets of axioms which characterize these structures (up to isomorphism).

For $\mathbb{R}$ as a linearly ordered set how do you handle Cauchy completeness or Dedekind completeness, and is one easier than the other?

Of course, your axioms should be first order.

1.7 Can you characterize, in a first order way, the class of metric spaces?

What about the class of topological spaces?

1.8 Consider a relational language, with no constants symbols and no operation symbols. For such a language we should also allow the empty structure (that structure on the empty set). For traditional reasons that doesn’t happen. (This rarely, if ever, causes a problem in model theory, but can in other parts of mathematics.)

Can you think of an explanation of how this tradition might have come about?

1.2 Basic notions

In this section we gather together a few more basic notions and set up some conventions which make the day to day handling of formulas a little less tedious.

We begin with the simplest comparison between two structures.

1.10 DEFINITION. Given a pair $\mathfrak{A}, \mathfrak{B}$ of structures (for the same language), we write

$$\mathfrak{A} \subseteq \mathfrak{B}$$

and say $\mathfrak{A}$ is a substructure of $\mathfrak{B}$ or $\mathfrak{B}$ is a superstructure of $\mathfrak{A}$, if the following hold.

- The carrier $A$ of $\mathfrak{A}$ is a subset of the carrier $B$ of $\mathfrak{B}$.

- For each constant symbol $K$ of the signature, $\mathfrak{A}[K] = \mathfrak{B}[K]$.

- For each $n$-placed relation symbol $R$ of the signature, the relation $\mathfrak{A}[R]$ is the restriction to $A$ of the relation $\mathfrak{B}[R]$. In other words we have

$$\mathfrak{A}[R]a_1 \cdots a_n \iff \mathfrak{B}[R]a_1 \cdots a_n$$

for each $a_1, \ldots, a_n \in A$. 
1. Syntax and semantics

• For each \(n\)-placed relation symbol \(O\) of the signature, the set \(A\) is closed under \(\mathcal{B}[O]\), and the operation \(\mathcal{A}[O]\) is the restriction to \(A\) of the operation \(\mathcal{B}[O]\). In other words we have

\[
\mathcal{A}[O]a_1 \cdots a_n = \mathcal{B}[O]a_1 \cdots a_n
\]

for each \(a_1, \ldots, a_n \in A\).

In short, \(A\) is closed under the attributes of \(\mathcal{B}\), and these give the attributes of \(\mathcal{A}\). ■

Thus, given a structure \(\mathcal{B}\), each substructure \(\mathcal{A}\) is completely determined by its carrier. However, not every subset of the carrier of \(\mathcal{B}\) is the carrier of a substructure.

In Section 3.1 we will look at a generalization of this idea. We define the notion of an embedding

\[
\mathcal{A} \xrightarrow{f} \mathcal{B}
\]

of a structure into another using a function \(f\) between the carriers. When this function is an insertion we obtain \(\mathcal{A} \subseteq \mathcal{B}\).

Two structures \(\mathcal{A}, \mathcal{B}\) (of the same language) are isomorphic

\[
\mathcal{A} \cong \mathcal{B}
\]

if there is an isomorphism from one to the other. This is a bijection between the carriers which matches the distinguished attributes. We needn’t write down the precise details of this, for it is a particular case of an embedding (which we look at later). However, we will use the notion.

Each language \(L\) consists of a set of formulas, some of which are sentences. Often we are interested in particular kinds of formulas, ones of a certain ‘syntactic complexity’. The most common measure of formulas is by quantifier complexity.

There are various useful classifications of formulas. Let’s look at two of these, one of which builds on top of the other.

Let \(L\) be an arbitrary language.

• An atom is just an atomic formula, as in Definition 1.6.

• A literal is an atom \(\alpha\) or the negation \(\neg \alpha\) of an atom.

• A formula \(\delta\) is quantifier-free if it contains no quantifiers, no uses of \(\forall\) or \(\exists\).

Each quantifier-free formula \(\delta\) can be rephrased in one of two useful normal forms.

• The conjunctive normal form in which \(\delta\) is rephrased as a conjunction

\[
D_1 \land \cdots \land D_m
\]

of disjuncts each of which is a disjunction of literals.

• The disjunctive normal form in which \(\delta\) is rephrased as a disjunction

\[
C_1 \lor \cdots \lor C_m
\]

of conjuncts each of which is a conjunction of literals.
Here ‘rephrased’ means ‘is logically equivalent to’. The standard boolean manipulation of formulas enables us to move from $\delta$ to either of the normal forms.

In the same way, using the rules for manipulating quantifiers, we know that each formula can be rephrased in prenex normal form as

$$(Q_1 v_1) \cdots (Q_n v_n) \delta$$

where each $Q$ in the prenex is a quantifier and the matrix $\delta$ is quantifier-free. By taking note of the alternations in the prenex we obtain the quantifier hierarchy of formulas.

Thus $\forall_0 = \exists_0 = QF$ is the set of formulas each of which is logically equivalent to a quantifier-free formula. For each $n \in \mathbb{N}$ the sets

$$\forall_{n+1} \quad \exists_{n+1}$$

consists of the formulas logically equivalent to

$$(\forall v_1, \ldots, v_n) \phi \quad (\exists v_1, \ldots, v_n) \phi$$

where

$$\phi \in \exists_n \quad \phi \in \forall_n$$

respectively. Inclusions between these sets are indicated in the diagram.

As you can imagine, handling formulas can be a bit tedious especially if we stick strictly to the letter of the law.

When we display a particular formula we often leave out some of the brackets or use different shapes of brackets to make the formula more readable. There are one or two other tricks we sometimes use.

Each formula $\phi$ has an associated set $\partial \phi$ of free variables. We often write

$$\phi(v_1, \ldots, v_n)$$

to indicate that $\partial \phi = \{v_1, \ldots, v_n\}$. Notice that at this level we are not concerned with the order or number of occurrences of each free variable in $\phi$. In fact, a variable can occur freely and bound in the same formula. Exercise 1.1 gives an extreme example of what can happen.

For much of what we do often a finite list of variables can be treated as a single variable. We use some informal conventions to handle this. Thus we often write

$$\phi(v)$$

to indicate that $v$ is a list

$$v_1, \ldots, v_n$$
of variables each of which occurs freely in $\phi$. Thus, in this usage

$$\phi(v) \quad \phi(v_1, \ldots, v_n)$$

mean the same thing. We sometimes refer to such a list $v$ as a batch of formulas. There is, of course, plenty of scope for confusion here. However, we will always make the situation clear.

In the same way we sometimes write

$$(\forall v)\phi(v) \quad \text{for} \quad (\forall v_1, \ldots, v_n)\phi(v_1, \ldots, v_n)$$

when the number of variables in the batch $v$ is not important.

As well as single formulas we also use sets of formulas. Let $\Gamma$ be such a set of formulas, and consider

$$\partial \Gamma = \bigcup \{ \partial \phi \mid \phi \in \Gamma \}$$

the set of all variables that occur free somewhere in $\Gamma$. This could be an infinite set. Often we restrict this support.

1.11 **DEFINITION.** A type is a set $\Gamma$ of formulas such that $\partial \Gamma$ is finite. ■

Do not confuse this use of the word ‘type’ with other uses. Some of these have no relationship at all with this usage. You are also warned that the word ‘type’ is sometimes used in model theory in a more restrictive sense. We will have more to say about this later.

**Exercises**

1.9 Consider the following three posets.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>carrier</td>
<td>${a, b}$</td>
<td>${a, b, c}$</td>
<td>${a, b, c, d}$</td>
</tr>
<tr>
<td>comparisons</td>
<td>$a \leq b$</td>
<td>$a \leq b \leq c$</td>
<td>$d \leq a \leq c, ; d \leq b \leq c$</td>
</tr>
</tbody>
</table>

Draw a picture of each and determine those pairs of structures where one is a substructure of the other.

1.10 Consider a relational language (just one 2-placed relation will do). Write down the definition of an isomorphism between such structures.

1.11 In the exercises for Section 1.1 you wrote down axioms for various classes of structures.

Determine the quantifier complexity of each set.

Can you spot something?

1.12 In an arbitrary language consider just the quantifier-free formulas. (You may use the propositional language.)

Describe two algorithms which consume an arbitrary formula and return the conjunctive normal form and the disjunctive normal form of the formula.

What about an algorithm that converts one normal form into the other?

Can you see any practical use of these normal forms?

What about the prenex normal form of an arbitrary formula?
1.3 Satisfaction

Having made the effort to set up the notion of the language $L$ suitable for structures $\mathfrak{A}$ of some given signature, and in which the idea of a sentence $\sigma$ is made precise, it is now patently obvious what

the structure $\mathfrak{A}$ satisfies $\sigma$

means. The whole of the rather tedious construction of $L$ was driven with this in mind. Nevertheless, it is worth looking at the formal definition of this notion (not least because some people think that it has some content). In more advanced work various non-elementary languages are used, and then the internal workings of the language and its satisfaction relation are more important.

We are talking about the pivotal notion of model theory, so we can’t keep writing it out in words. Accordingly we let

$$\mathfrak{A} \models \sigma$$

abbreviate the phrase above ($\mathfrak{A}$ satisfies $\sigma$). This, of course, is a relation between structures $\mathfrak{A}$ and sentences $\sigma$, so each instance is either true or false. This truth value is generated by recursion on the construction of $\sigma$. That is, we define outright the value for simple sentences, and then show how to obtain the value for compound sentences from the values for its components. This brings out a minor snag.

As we unravel the construction of $\sigma$, we will meet certain formulas, and these may contain free variables. But such variables have no interpretation in $\mathfrak{A}$. They are merely a linguistic device to indicate certain bindings in larger formulas. However, to push through the recursion, we need to show what to do with any free variables that arise as we unravel the sentence $\sigma$. We use a little trick.

1.12 DEFINITION. For a structure $\mathfrak{A}$, an $\mathfrak{A}$-assignment is a function $x$ which attaches to each variable $v$ an element $vx$ of $\mathfrak{A}$. ■

Notice that although we call $x$ a function, we write its input $v$ on the left.

The idea is that if we meet a free variable $v$ which, apparently, has no interpretation, then we give it the value $vx$. Before we see how this helps with the satisfaction relation, let’s look at a similar use in a simpler situation.

Each term $t$ (of the underlying language $L$) is built from certain constants $K$, certain operation symbols $O$, and certain free variables $v$. How can we give $t$ a value in some structure $\mathfrak{A}$? We have interpretations $\mathfrak{A}[K]$ and $\mathfrak{A}[O]$ of $K$ and $O$, but we have no interpretation of $v$. To get round this we use an assignment $x$, and define the value of $t$ in $\mathfrak{A}$ at $x$.

1.13 DEFINITION. For each structure $\mathfrak{A}$ (suitable for a language $L$), each $\mathfrak{A}$-assignment $x$, and each term $t$ (of $L$) the element

$$\mathfrak{A}[t]x$$

of $\mathfrak{A}$ is generated by recursion on the construction of $t$ using the following clauses.

- If $t$ is a variable $v$ then

$$\mathfrak{A}[t]x = vx$$

using the element assigned to $v$ by $x$. 
1. Syntax and semantics

- If \( t \) is a constant symbol \( K \) then
  \[
  \mathfrak{A}[t]x = \mathfrak{A}[K]
  \]
  (the interpretation of \( K \) in \( \mathfrak{A} \)).

- If \( t \) is a compound
  \[(O\ t_1 \ldots t_n)\]
  where \( O \) is an \( n \)-placed operation symbol and \( t_1, \ldots, t_n \) are smaller terms, then
  \[
  \mathfrak{A}[t]x = \mathfrak{A}[O]a_1 \cdots a_n
  \]
  where
  \[
  a_i = \mathfrak{A}[t_i]x
  \]
  for each \( 1 \leq i \leq n \).

No other clause are required.

There is nothing in this definition. Each term \( t \) names, in an obvious way, a certain (compound) operation on (the carrier of) \( \mathfrak{A} \). This construction merely evaluates this operation, in the obvious way, where the inputs are supplied by the assignment \( x \). In particular, for each term \( t \) almost all of the assignment \( x \) is not needed.

1.14 LEMMA. Let \( \mathfrak{A} \) be a structure and let \( t \) be a term (of the underlying language). If \( x \) and \( y \) are \( \mathfrak{A} \)-assignments which agree on \( \partial t \), that is if
  \[
  \nu x = \nu y
  \]
  for each \( \nu \in \partial t \), then
  \[
  \mathfrak{A}[t]x = \mathfrak{A}[t]y
  \]
  holds.

This is proved by the obvious induction over the construction of \( t \).

To generate the satisfaction relation we use the same trick. We define a more general relation
  \[
  \mathfrak{A} \models \phi x
  \]
which says
  \( \mathfrak{A} \) satisfies the formula \( \phi \) where each free variable \( \nu \) takes the value \( \nu x \)
and then in Lemma 1.16 we show the irrelevancy of most of \( x \).

1.15 DEFINITION. For each structure \( \mathfrak{A} \) (suitable for a language \( L \)), each \( \mathfrak{A} \)-assignment \( x \), and each formula \( \phi \) (of \( L \)) the truth value
  \[
  \mathfrak{A} \models \phi x
  \]
is generated by recursion on the structure of \( t \) using the following clauses.

\[
\begin{align*}
\mathfrak{A} \models (\text{true})x & \iff \text{true} \\
\mathfrak{A} \models (\text{false})x & \iff \text{false} \\
\mathfrak{A} \models (t_1 \equiv t_2)x & \iff \mathfrak{A}[t_1]x = \mathfrak{A}[t_2]x \\
\mathfrak{A} \models (Rt_1 \cdots t_n)x & \iff \mathfrak{A}[R]a_1 \cdots a_n \text{ where } a_i = \mathfrak{A}[t_i]x \\
\mathfrak{A} \models (\neg \psi)x & \iff \neg (\mathfrak{A} \models \psi)x \\
\mathfrak{A} \models (\theta \land \psi)x & \iff \mathfrak{A} \models \theta x \text{ and } \mathfrak{A} \models \psi x \\
\mathfrak{A} \models (\theta \lor \psi)x & \iff \mathfrak{A} \models \theta x \text{ or } \mathfrak{A} \models \psi x \\
\mathfrak{A} \models (\theta \rightarrow \psi)x & \iff \neg (\mathfrak{A} \models \theta x) \text{ or } \mathfrak{A} \models \psi x \\
\mathfrak{A} \models ((\forall v)\psi)x & \iff \mathfrak{A} \models \psi y \quad \text{for each } \mathfrak{A}\text{-assignment } y \\
\mathfrak{A} \models ((\exists v)\psi)x & \iff \mathfrak{A} \models \psi y \quad \text{for some } \mathfrak{A}\text{-assignment } y
\end{align*}
\]

No other clauses are required.

Notice that
\[
\mathfrak{A} \models (\text{true})x
\]
always holds, whereas
\[
\mathfrak{A} \models (\text{false})x
\]
ever holds. This is the principal job of these two constant sentences.

To get the original satisfaction relation (for sentences) we make the following observation.

1.16 LEMMA. Let \( \mathfrak{A} \) be a structure and let \( \phi \) be a formula (of the underlying language). If \( x \) and \( y \) are \( \mathfrak{A} \)-assignments which agree on \( \partial \phi \), that is if
\[
v_x = v_y
\]
for each \( v \in \partial \phi \), then
\[
\mathfrak{A} \models \phi x \iff \mathfrak{A} \models \phi y
\]
holds.

By definition, a sentence is a formula \( \sigma \) with \( \partial \sigma = \emptyset \). Vacuously, for such a sentence, each two \( \mathfrak{A} \)-assignments \( x \) and \( y \) agree on \( \partial \sigma \), and hence
\[
\mathfrak{A} \models \sigma x \iff \mathfrak{A} \models \sigma y
\]
holds. In other words, either
\[
\mathfrak{A} \models \sigma x
\]
for every \( \mathfrak{A} \)-assignment or for no \( \mathfrak{A} \)-assignment. Thus, we may write
\[
\mathfrak{A} \models \sigma
\]
to indicate that $\mathfrak{A} \models \sigma x$ holds for every $x$.

In a similar way we may simplify the notation $\mathfrak{A} \models \phi x$.

Consider a formula $\phi(v_1, \ldots, v_n)$, that is a formula $\phi$ with $\partial \phi = \{ v_1, \ldots, v_n \}$. Given a structure $\mathfrak{A}$ and an assignment $x$, the truth value of

$$\mathfrak{A} \models \phi x$$

depends only on the elements

$$a_1 = v_1 x, \ldots, a_n = v_n x$$

selected from $x$ by the free variables. Thus we often write

$$\mathfrak{A} \models \phi(a_1, \ldots, a_n)$$

in place of the official notation.

Sometimes we go even further. By a point $a$ of the structure $\mathfrak{A}$ we mean a list $a_1, \ldots, a_n$ of elements of $\mathfrak{A}$. We may then write

$$\mathfrak{A} \models \phi(a)$$

for the satisfaction relation. Of course, this assumes there is a match between the point $a$ and the batch $v$ of free variables of $\phi$.

In Section 1.2 we introduced the idea of isomorphic structures

$$\mathfrak{A} \simeq \mathfrak{B}$$

(of the same signature). There is a semantic analogue of this.

1.17 DEFINITION. Let $\mathfrak{A}, \mathfrak{B}$ be structures of the same signature.

We write

$$\mathfrak{A} \equiv \mathfrak{B}$$

and say $\mathfrak{A}$ and $\mathfrak{B}$ are elementarily equivalent if

$$\mathfrak{A} \models \sigma \iff \mathfrak{B} \models \sigma$$

holds for each sentence $\sigma$ (of the underlying language).

Almost trivially

$$\mathfrak{A} \simeq \mathfrak{B} \implies \mathfrak{A} \equiv \mathfrak{B}$$

holds (but the proof of this is rather tedious). However, the converse is false (in general). In fact, as we will see in Theorem 1.28, it can happen that $\mathfrak{A} \equiv \mathfrak{B}$ but $\vert \mathfrak{A} \vert \neq \vert \mathfrak{B} \vert$ and so these structures can’t be isomorphic.
1.3. Satisfaction

Exercises

1.13 Sketch the proofs of Lemmas 1.14 and 1.16.

1.14 Consider the formulas $\phi_r$ generated in Exercise 1.1, and let $\mathcal{N} = (\mathbb{N}, <)$.

(a) Characterize the elements $a$ of $\mathcal{N}$ such that $\mathcal{N} \models \phi_r(a)$.

(b) Show that we have

$$\mathcal{N} \models (\forall v)[\phi_{r+1} \rightarrow \phi_r]$$

for each $r < \omega$.

(c) Describe the formulas $\psi(v)$ for which we have

$$\mathcal{N} \models (\forall v) [\psi \rightarrow \phi_r]$$

for all $r < \omega$.

1.15 Consider the language on the empty signature. In other words, consider the language of pure equality.

(a) Show that for each $n < \omega$ there are sentences $\sigma_n$ and $\tau_n$ such that

$$\mathcal{A} \models \sigma_n \iff |\mathcal{A}| \geq n \quad \mathcal{A} \models \tau_n \iff |\mathcal{A}| = n$$

for each structure $\mathcal{A}$. What is the quantifier complexity of each of these sentences?

(b) Show there is a set $\text{Inf}$ of sentences such that

$$\mathcal{A} \models \text{Inf} \iff \mathcal{A} \text{ is infinite}$$

for each structure $\mathcal{A}$. Is there a finite set of such sentences?

(c) Is there a set $\text{Fin}$ of sentences such that

$$\mathcal{A} \models \text{Fin} \iff \mathcal{A} \text{ is finite}$$

for each each structure $\mathcal{A}$?

1.16 Consider the signature with just of symbol which is a 1-placed operation symbol. For that language produce a sentence $\sigma$ such that we have

$$\mathcal{A} \models \sigma \implies \mathcal{A} \text{ is infinite}$$

for each structure $\mathcal{A}$ of that signature.

What is the quantifier complexity of your sentence? Is there such a sentence of lower quantifier complexity? (If there is, why didn’t you use that one?)

Can you strengthen this to an equivalence?

1.17 Suppose $\mathcal{A} \subseteq \mathcal{B}$.

Show that

$$\mathcal{A} \models \delta(a) \iff \mathcal{B} \models \delta(a)$$

for each quantifier-free formula $\delta(v)$ and point $a$ of $\mathcal{A}$ which matches the batch $v$ of free variables of $\delta$.

Show that

$$\mathcal{A} \models \theta(a) \implies \mathcal{B} \models \theta(a)$$

for each $\exists_1$-formula $\theta(v)$ and point $a$ of $\mathcal{A}$ which matches the batch $v$ of free variables of $\theta$.

Find an example to show that this implication is not an equivalence.
1.4 Consequence

Each language $L$ has an associated satisfaction relation

$$\mathfrak{A} \models \sigma$$

between structures $\mathfrak{A}$ (for $L$) and sentences $\sigma$ (of $L$). We can refine this.

1.18 DEFINITION. Let $L$ be a language, let $\mathcal{K}$ be a class of structures for $L$, and let $\Sigma$ be a set of sentences of $L$.

(a) The relation

$$\mathcal{K} \models \Sigma$$

holds if

$$\mathfrak{A} \models \sigma$$

for each $\mathfrak{A} \in \mathcal{K}$ and $\sigma \in \Sigma$.

(b) The theory $\text{Th}(\mathcal{K})$ of $\mathcal{K}$ is the set of all sentences $\sigma$ such that $\mathcal{K} \models \sigma$.

(c) The class $\text{Md}(\Sigma)$ of models of $\Sigma$ is the class of all structures $\mathfrak{A}$ such that $\mathfrak{A} \models \Sigma$.

The two assignments $\text{Md}(\cdot)$ and $\text{Th}(\cdot)$ form a galois connection. Thus we have

$$\mathcal{K} \subseteq \text{Md}(\Sigma) \iff \Sigma \subseteq \text{Th}(\mathcal{K})$$

for each class $\mathcal{K}$ (of structures) and each set $\Sigma$ (of sentences). Furthermore, either side holds precisely when

$$\mathcal{K} \models \Sigma$$

holds. In particular, both the composites

$$\text{Th} \circ \text{Md} \quad \text{Md} \circ \text{Th}$$

are closure operations and, as expected, we look at the closed families.

1.19 DEFINITION. (a) A set $T$ of sentences is a theory if $T = \text{Th}(\mathcal{K})$ for some class $\mathcal{K}$ of structures. Equivalently, $T$ is a theory if (and only if) $T = \text{Th}(\text{Md}(T))$.

(b) A class $\mathcal{K}$ of structures is elementary or an elementary class if $\mathcal{K} = \text{Md}(\Sigma)$ for some set $\Sigma$ of sentence. Equivalently, $\mathcal{K}$ is elementary if (and only if) $\mathcal{K} = \text{Md}(\text{Th}(\mathcal{K}))$.

These notions prompt some obvious questions.

- Are there any necessary and sufficient conditions for a class to be elementary?
- By definition, a class is elementary if it has the form $\text{Md}(\Sigma)$ for a set $\Sigma$ of sentences. When is a class strictly elementary, that is when does it have the form $\text{Md}(\sigma)$ for a single sentence?
- What are the necessary and sufficient conditions for a set of sentences to be a theory?
- By definition, a set $T$ is a theory if it has the form $\text{Th}(\text{Md}(\Sigma))$ for some set $\Sigma$. We then say $\Sigma$ axiomatizes $T$. When does a theory have a finite set of axioms? When is a theory $\forall_n$-axiomatizable for some $n$?
To answer these and similar questions we need a tool, the compactness theorem. This is the pivotal method of model theory, and is discussed in detail in the next section. Here we see how it relates to another part of mathematical logic.

1.20 **DEFINITION.** A set $\Sigma$ of sentences (of some language) is **consistent** or **satisfiable** if it has a model, that is if $\mathfrak{A} \models \Sigma$ for some structure $\mathfrak{A}$.

The theory $T$ is inconsistent (not consistent) if $T = \text{Th}(\emptyset)$, in which case $T$ is the set of all sentences of the language. Rarely do we need to consider this theory, so we often say ‘a theory $T$’ when we mean ‘a consistent theory $T$’.

1.21 **DEFINITION.** A theory $T$ is **complete** if it is consistent and $\mathfrak{A} \equiv \mathfrak{B}$ for all models $\mathfrak{A}, \mathfrak{B}$ of $T$.

A theory $T$ is $\kappa$-categorical (for a cardinal $\kappa$) if $\mathfrak{A} \cong \mathfrak{B}$ for all models $\mathfrak{A}, \mathfrak{B}$ of $T$ with $|\mathfrak{A}| = \kappa = |\mathfrak{B}|$.

It is an easy exercise to see that a theory is complete if and only if it has the form $\text{Th}(\mathfrak{A})$ for a structure $\mathfrak{A}$. Another easy exercise (but using a result we haven’t yet seen) shows that if a consistent theory $T$ in a language $L$ is $\kappa$-categorical for some $\kappa \geq |L|$, then it is complete. Roughly speaking a complete theory is a large theory in the sense that it can’t take in any more sentences without becoming inconsistent. At the other extreme, the pure logic of a language is the theory of the class of all structures for that language. Thus a sentence belongs to this theory if and only if it is universally valid.

1.22 **DEFINITION.** For a set $\Sigma$ of sentences and a sentence $\sigma$ we write

$$\Sigma \vdash \sigma$$

and say $\Sigma$ **entails** $\sigma$ or $\sigma$ is a **consequence** of $\Sigma$ if $\sigma \in \text{Th}(\Sigma)$, that is if $\mathfrak{A} \models \sigma$ for each model $\mathfrak{A} \models \Sigma$.

In this notation $\Sigma$ is a set of axioms for a theory $T$ exactly when we have

$$\sigma \in T \iff \Sigma \vdash \sigma$$

for each sentence $\sigma$. Often we describe a theory by writing down a particular set of axioms. Then one of the problems is to characterize all (or a large amount of) the consequences of the axioms. At other times the problem can be to axiomatize the theory of some given class of structures (which is described in a non-elementary way).

Although it is not strictly part of model theory, at this point it is worth comparing this semantic consequence relation with the proof theoretic consequence relation, the kind of consequence relation met in a course on predicate calculus (and analysed in more detail in Proof Theory).

The Definition 1.22 of the relation $\vdash$ involves an external quantification over a potentially large class of structures. The relation

$$\Sigma \vdash \sigma$$
holds if

\[ \text{for all structures } \mathcal{A} \text{ we have } \ldots \]

where \ldots describes an interaction between \( \mathcal{A}, \Sigma, \text{ and } \sigma \). However, \( \Sigma \vdash \sigma \) is a relation between syntactic objects, and the question arises of whether it can be characterized in purely combinatorial terms.

Gödel’s \textbf{completeness theorem} shows that it can.

(You should not confuse the two different uses of ‘completeness’ here. They are related, but not the same.)

To analyse \( \vdash \) we first set up a proof-theoretic relation

\[ \Sigma \mid\sim \sigma \]

between sets \( \Sigma \) of sentences and sentences \( \sigma \). The essential feature of this is that it is entirely combinatorial. This relation holds if and only if there is a certain (finite) configuration of strings of symbols. The intended semantics is not referred to at all. Thus, \( \Sigma \mid\sim \sigma \) can be shown to hold by exhibiting a certain collection of symbols formatted in an appropriate way.

(The notation \( \mid\sim \) is not standard. Often \( \vdash \) is used for both the semantic consequence and the proof theoretic consequence relations. Here we want to compare the two, so it is convenient to have different notations. Proof theoretic consequence is not used again after this chapter, except for a mention in Chapter [**Construction Technique**].)

There are several different ways of setting up \( \mid\sim \), most of which are needed for one job or another (and some of which are entirely untainted by content and interest). Here we needn’t worry about the precise details.

The analysis now investigates the relationship between

\[ \mid\sim \vdash \]

to produce two particular results, one minor and one major.

- The relation \( \mid\sim \) is \textbf{sound}, that is

\[ \Sigma \mid\sim \sigma \Rightarrow \Sigma \vdash \sigma \]

for all \( \Sigma \) and \( \sigma \). This is a relatively trivial observation.

- The relation \( \mid\sim \) is \textbf{adequate}, that is

\[ \Sigma \vdash \sigma \Rightarrow \Sigma \mid\sim \sigma \]

for all \( \Sigma \) and \( \sigma \). This requires quite a bit of work.

- The combination of these two results is the \textbf{completeness theorem}, that is

\[ \Sigma \vdash \sigma \iff \Sigma \mid\sim \sigma \]

for all \( \Sigma \) and \( \sigma \).
Because of the way $\neg$ is set up we observe that if

$$\Sigma \vdash \sigma$$

then

$$\Gamma \vdash \sigma$$

for some finite part $\Gamma$ of $\Sigma$. This leads to the following result.

**1.23 THEOREM.** Let $\Sigma$ be a set of sentences (in some language). If each finite part of $\Sigma$ has a model, then $\Sigma$ has a model.

**Proof.** We prove the contrapositive. Thus suppose $\Sigma$ does not have a model. Then, vacuously we have

$$\Sigma \vdash \sigma$$

for each sentence $\sigma$. Consider the sentence false which does not have a model. We have

$$\Sigma \vdash \text{false}$$

and hence

$$\Sigma \vdash \text{false}$$

by the adequacy of $\neg$. But now

$$\Gamma \vdash \text{false}$$

for some finite part $\Gamma$ of $\Sigma$, and then

$$\Gamma \vdash \text{false}$$

by the soundness of $\neg$. This shows that $\Gamma$ does not have a model. \[\square\]

This result is a version of the compactness theorem.

**Exercises**

The first few of the this collection of exercises are concerned with capturing certain notions in a first order fashion. You may want to have another look at the exercises for Section 1.1.

**1.18** An equivalence structure has the form $(A, R)$ where $R$ is an equivalence relation on the carrier $A$.

(a) Axiomatize the class of equivalence structures.

(b) For $m, n < \omega$, axiomatize the class of equivalence structures each having no more than $m$ blocks (equivalence classes), where each of these has no more than $n$ members.

(c) Axiomatize the class of equivalence structures having infinitely many equivalence classes where each of these is infinite.

(d) Axiomatize the class of equivalence structures which are such that if there is a finite equivalence class, then there is an equivalence class of each larger finite size.
1.19 (a) Axiomatize the classes of posets, linear orderings, dense linear orderings, and discrete linear orderings.
(b) Write down formulas \( \theta(u, v, w) \), \( \psi(u, v, w) \), \( \phi(u, v) \) such that
\[
\mathfrak{A} \models \theta(a, b, c) \iff c \text{ is the l.u.b. of } b \text{ and } c
\]
\[
\mathfrak{A} \models \psi(a, b, c) \iff a, b, c \text{ are linearly ordered in } \mathfrak{A}
\]
if \( a, b \) are comparable and not equal, then exactly
\[
\mathfrak{A} \models \phi(a, b) \iff \text{three elements lie strictly between } a, b, \text{ and these}
\]
three elements are pairwise incomparable
for each poset \( \mathfrak{A} \) and elements \( a, b, c \) of \( \mathfrak{A} \).

The next exercise uses the language suitable for structures
\[
\mathfrak{A} = (A, *, e)
\]
where * is a binary operation on \( A \) and \( e \) is a distinguished element.

1.20 (a) Axiomatize the classes of groups, abelian groups, torsion-free abelian groups, and divisible abelian groups. Which of these classes are finitely axiomatizable?
Write down formulas \( \theta(u), \psi(u, v), \phi(u, v) \) such that
\[
\mathfrak{A} \models \theta(a) \iff a \text{ is a commutator}
\]
\[
\mathfrak{A} \models \psi(a, b) \iff a \text{ is in the centralizer of } b
\]
\[
\mathfrak{A} \models \phi(a, b) \iff \text{there is an inner automorphism taking } a \text{ to } b
\]
for each group \( \mathfrak{A} \) and elements \( a, b \) of \( \mathfrak{A} \).

1.21 By using a suitable signature axiomatize the classes of rings (with 1), commutative rings (with 1), integral domains, integral domains of characteristic \( p \) (where \( p \) is a prime), integral domains of characteristic 0, fields, algebraically closed fields.
Which of the classes are \( \forall_2 \)-axiomatizable, \( \forall_1 \)-axiomatizable, finitely axiomatizable?

1.22 Consider the reals as a structure \((\mathbb{R}, +, \times, \leq, 0, 1)\) (with the obvious attributes). Look up the axioms which characterize this structure up to isomorphism. Observe that most of these are first order, but the crucial one isn’t. What is this non-elementary axiom?

1.23 Let \( R \) be a ring with 1, and consider the right \( R \)-modules. Think of each of these as a structure
\[
\mathfrak{A} = (A, +, 0, (f_r \mid r \in R))
\]
where \((A, +, 0)\) is an abelian group and, for each \( r \in R \), the 1-placed operation \( f_r \) is \( a \mapsto ar \). Thus \( R \) is used to index part of the signature.
Write down axioms for this class of modules.
What changes need to be made to axiomatize the class of left \( R \)-modules?

1.24 Show that for each consistent theory \( T \) the following are equivalent.
(i) \( T \) is complete
(ii) For each sentence \( \sigma \), if \( T \cup \{\sigma\} \) is consistent, then \( \sigma \in T \).
(iii) For each theory \( T' \), if \( T \subseteq T' \) then either \( T = T' \) or \( T' \) is inconsistent.
(iv) For each pair \( \sigma, \tau \) of sentences, if \( \sigma \lor \tau \in T \), then \( \sigma \in T \) or \( \tau \in T \).
(v) For each sentence \( \sigma \) either \( \sigma \in T \) or \( \neg \sigma \in T \).
(vi) There is a structure \( \mathfrak{A} \) with \( T = Th(\mathfrak{A}) \).
1.5. Compactness

The following exercise is quite tricky, but its solution uses an important technique which you should learn as soon as possible.

1.25 Show that if no finite extension of a (consistent) theory is complete, then the theory has at least \(2^{\aleph_0}\) complete extensions.

Finally, here is almost all you need to know about galois connections. (Be careful. Some of the ignoscenti confuse galois connections with poset adjunctions.)

1.26 Let \(A, S\) be a pair of posets with elements \(a, b, c, \ldots\) and \(r, s, t, \ldots\), respectively. Let

\[
\begin{array}{c}
A \xrightarrow{\mathfrak{r}} S \\
A \xleftarrow{\mathfrak{r}} S
\end{array}
\]

be a pair of assignments such that

\[a \leq \mathfrak{r}(s) \iff s \leq \mathfrak{r}(a)\]

holds for each \(a \in A\) and \(s \in S\).

(a) Show that both the composites \(\mathfrak{r} \circ \mathfrak{r}\) and \(\mathfrak{r} \circ \mathfrak{r}\) are inflationary.

(b) Show that \(\mathfrak{r} \circ \mathfrak{r} \circ \mathfrak{r} = \mathfrak{r}\) and \(\mathfrak{r} \circ \mathfrak{r} \circ \mathfrak{r} = \mathfrak{r}\) hold.

(c) Show that both \(\mathfrak{r}\) and \(\mathfrak{r}\) are antitone.

(d) Show that both \(\mathfrak{r} \circ \mathfrak{r}\) and \(\mathfrak{r} \circ \mathfrak{r}\) are closure operations.

1.5 Compactness

At the end of the previous section we obtained a statement (and an indication of one proof) of the compactness theorem. In this section we take a closer look at this result.

1.24 DEFINITION. A set \(\Sigma\) of sentences (of some language) is consistent or satisfiable if it has a model, that is if \(\mathfrak{A} \models \Sigma\) for some structure \(\mathfrak{A}\).

A set \(\Sigma\) of sentences is finitely satisfiable if each finite part has a model.

In this terminology, we saw that the completeness theorem implies the following.

1.25 THEOREM. (The crude compactness theorem) If a set of sentences (of some language) is finitely satisfiable, then it is satisfiable.

How should we prove this?

We have seen already one method of proof. We set up a proof-theoretic consequence relation \(\models\) and then prove a completeness result. The compactness result is an immediate consequence. However, this is not entirely satisfactory, for two reasons.

Firstly, we must set up all the machinery for \(\models\), and this takes some time. Furthermore, this machinery is never used again in model theory. (It may be used elsewhere in mathematical logic, but then \(\models\) will be the principal object of study, and it will be designed with some specific class of tasks in mind.)

Secondly, at the heart of the proof of completeness a certain structure is constructed. The method of construction can be modified to give a direct proof of compactness without a detour through \(\models\) and its properties.
The **witnessing construction** is described in Chapter [secC10]. This is a method of producing a structure out of a certain kind of family of sets of sentences (called a **consistency property**). In the first instance this construction gives us both compactness and completeness, virtually by the same proof. This method of construction is quite flexible, and gives us quite a lot of control over the end product. This is used to advantage in more advanced work. Furthermore, the same method can be lifted to higher order languages (but, of course, this requires a bit more work).

Another idea on how to prove compactness should have occurred to you.

Let $\Sigma$ be a finitely satisfiable set of sentences, and let $\Delta$ be the set of finite subset $\Delta$ of $\Sigma$. We are given a model $A(\Delta)$ of each such $\Delta \in \Delta$. Is there a way of patching together, in a coherent fashion, all of these $A(\Delta)$ to produce a model of $\Sigma$? There is, and it is called the **ultraproduct construction**. This is described in Chapter [uprod].

For the time being we do not need the details of the proof of Theorem 1.25, so let’s look at some applications of compactness.

1.26 THEOREM. Let $L$ be any language.

(a) The class of all infinite structures (for $L$) is elementary but not strictly elementary.

(b) The class of all finite structures is not elementary.

(c) A sentence (of $L$) holds in all infinite structures if and only if it holds in all sufficiently large structures.

(d) The theory of the class of finite structures has an infinite model.

(e) The theory of the class of infinite structures has no finite model.

Proof. (a) By Exercise 1.15 we know that for each $n < \omega$ there is a sentence $\sigma_n$ such that

$$A \models \sigma_n \iff |A| \geq n$$

for each structure $A$. Let

$$Inf = \{\sigma_n \mid n < \omega\}$$

so that $Md(Inf)$ is exactly the class of infinite structures. In particular, this class is elementary.

By way of contradiction, suppose that this class is strictly elementary. Thus there is a single sentence $\tau$ such that

$$A \models \tau \iff A \text{ is infinite}$$

for each structure $A$. In particular,

$$A \models \neg \tau \iff A \text{ is finite}$$

for each structure $A$. We show that the set

$$Inf \cup \{\neg \tau\}$$

is consistent, which is the required contradiction.

Any finite subset of this set is a subset of

$$\{\sigma_0, \ldots, \sigma_n, \neg \tau\}$$
for some $n < \omega$. By considering a sufficiently large finite structure, we see that this subset has a model. Thus $\text{Inf} \cup \{\neg \tau\}$ is finitely satisfiable and hence, by the compactness property, is satisfiable, as required.

(b) By way of contradiction, suppose the class of finite structures is elementary. Thus there is a set $\text{Fin}$ of sentences such that

$$\mathfrak{A} \models \text{Fin} \iff \text{A is finite}$$

for each structure $\mathfrak{A}$. A slight modification of the argument used in (a) shows that the set

$$\text{Inf} \cup \text{Fin}$$

is finitely satisfiable, and hence is satisfiable. This is not so, since no structure is both infinite and finite.

(c) Suppose the sentence $\tau$ holds in all infinite structures. Then

$$\text{Inf} \vdash \tau$$

and hence, by compactness, we have

$$\sigma_0, \ldots, \sigma_n \vdash \tau$$

for some $n < \omega$. Thus $\tau$ holds in any structure $\mathfrak{A}$ with $|\mathfrak{A}| \geq n$.

Conversely, suppose there is some $n < \omega$ such that the sentence $\tau$ holds in each structure $\mathfrak{A}$ with $|\mathfrak{A}| \geq n$. Then

$$\sigma_n \vdash \tau$$

and hence

$$\text{Inf} \vdash \tau$$

to show that $\tau$ holds in all infinite structures.

(d) Let $T$ be the theory of the class of finite structures. For each $n < \omega$, any sufficiently large structure is a model of $T \cup \{\sigma_n\}$, and hence

$$T \cup \text{Inf}$$

is finitely satisfiable. By compactness, this set has a model, and hence $T$ has an infinite model.

(e) Now let $T$ be the theory of the class of infinite structure. Then $\text{Inf} \subseteq T$, and hence no finite structure can be a model of $T$. $\blacksquare$

Theorem 1.25 is the crude compactness result because it can be refined to extract more information. We use the cardinality $|L|$ of the underlying language.

1.27 THEOREM. (The refined compactness theorem) Let $\Sigma$ be a set of $L$-sentences (for some language). If $\Sigma$ is finitely satisfiable, then it has a model $\mathfrak{A}$ with $|\mathfrak{A}| = |L|$.

Notice that this refined version does not follow from completeness, as outline in the previous section. However, it is an immediate consequence of the witnessing construction, which allows us to control the size of the structure produced. This might not seem much, but it has some surprising consequences.
1.28 THEOREM. The theory \( \text{Th}(\mathbb{N}) \) of the natural numbers is not \( \aleph_0 \)-categorical. That is, there is a countable structure \( A \) with \( A \equiv \mathbb{N} \) and \( A \not\equiv \mathbb{N} \).

Proof. Notice that we didn't specify which language \( \text{Th}(\mathbb{N}) \) is formalized in. That is because it doesn't matter. The result holds no matter which language we used. However, it is useful to have numerals (constant terms) \( \lceil n \rceil \) in the language.

To prove the result we enrich the language by adding one new constant symbol \( a \), say. Look at the set

\[
\text{Th}(\mathbb{N}) \cup \{ (\lceil n \rceil \neq a) \mid n \in \mathbb{N} \}
\]

in this enriched language. This is finitely satisfiable. To see this notice that in any finite part there is a largest \( n \in \mathbb{N} \) such that \( \lceil n \rceil \) occurs, and then \((\mathbb{N}, n + 1)\) is a model.

By the refined compactness result, the set has a countable model \((A, a)\) where \( A \equiv \mathbb{N} \) and \( a \) is some distinguished element. There is a unique embedding \( \mathbb{N} \rightarrow A \) (given by \( n \mapsto A[\lceil n \rceil] \)) and the extra sentences ensure that \( a \) is not in the range of this. Thus \( A \not\equiv \mathbb{N} \).

This is sometimes known as Skolem's paradox (even though it is not a paradox, just a surprise). Skolem's original proof used a kind of ultraproduct construction. Later the trick used in this proof will be turned into a powerful tool.

Exercises

1.27 (a) Show that if \( \Sigma \vdash \tau \) (where \( \Sigma \) is a set of sentences and \( \tau \) is a sentence of the same language), then \( \Gamma \vdash \tau \) for some finite \( \Gamma \subseteq \Sigma \).

(b) Show that a consistent set \( \Sigma \) of sentences is a theory if and only if it contains all universally valid sentences and \( \tau \in \Sigma \) whenever \( \sigma, \sigma \rightarrow \tau \in \Sigma \).

1.28 Suppose

\[
\mathcal{K} = \bigcap \{ \mathcal{K}_r \mid r < \omega \}
\]

where \( \{ \mathcal{K}_r \mid r < \omega \} \) is a strictly descending chain of strictly elementary classes. Show that \( \mathcal{K} \) is elementary but not strictly elementary.

1.29 Let \( \mathcal{K} \) be a strictly elementary class (for some language) and suppose

\[
\mathcal{K} = \mathcal{L} \cup \mathcal{R} \quad \mathcal{L} \cap \mathcal{R} = \emptyset
\]

where both \( \mathcal{L} \) and \( \mathcal{R} \) are elementary. Show that both \( \mathcal{L} \) and \( \mathcal{R} \) are strictly elementary.

1.30 Let \( \mathcal{F}, \mathcal{F}_0, \mathcal{F}_p, \mathcal{F}_f \) be the classes of fields, fields of characteristic zero, fields of characteristic \( p \) (for a given prime \( p \)), a fields of finite (non-zero) characteristic, respectively. Let \( T, T_0, T_p, T_f \) be the respective theories of these classes.

(a) Which of these classes are elementary and which are strictly elementary.

(b) Which if these theories are finitely axiomatizable.

(c) Show that each sentence \( \tau \in T_0 \) holds in each field of sufficiently large (prime) characteristic.

(d) Show that \( T_f \) has a model of characteristic zero.
1.31 Let $\mathbb{R}$ be the real numbers viewed as a first order structure.
   Show there is a countable structure $\mathfrak{A}$ with $\mathfrak{A} \equiv \mathbb{R}$.
   Can you say what this structure might be?

1.32 In several solutions earlier in this chapter a use of the compactness theorem has been mentioned. You should now go back and fill in the details.
2
The effective elimination of quantifiers

Strictly speaking, the topic of this chapter, quantifier elimination, is not a part of model theory proper. It is included here for two reasons, one minor and one major. Anyone who claims to have some familiarity with mathematical logic should know something about quantifier elimination. That is the minor reason. The major reason is that the topic had a considerable influence on the early development of model theory, and we will follow an idealized version of that path. It could be said the quantifier elimination is a recurring theme throughout this book.

2.1 The generalities of quantifier elimination

Suppose $T$ is a theory in some language. It doesn’t matter how $T$ is describe. It could be given in the form $Th(\mathcal{K})$ for some class $\mathcal{K}$ of structures. It could be given as the consequences of some set of axioms. It could be given in some other way.

To understand $T$ we need to know something about the way quantification behaves in (the models of) $T$.

2.1 DEFINITION. Let $T$ be a theory in some language.
(a) Two formulas $\phi$ and $\psi$ are $T$-equivalent if

$$T \vdash (\forall v)[\phi \leftrightarrow \psi]$$

where $v$ is a batch of variables which includes $\partial \phi \cup \partial \psi$.

(b) The theory $T$ has $EQ$ (elimination of quantifiers) if each formula is $T$-equivalent to some quantifier-free formula.

Of course, if a theory has $EQ$ then it must be rather special. One of our long term aims (which we achieve in Chapter 4) is to characterize this speciality. In this chapter we begin with a few examples of this property.

How can we show that a theory $T$ has $EQ$? The obvious way is to describe an algorithm which, when supplied with a formula $\phi$, will return a quantifier-free formula $\psi$ which is $T$-equivalent to $\phi$. In this chapter we describe, in reasonable but not full detail, two examples of such an algorithm. We will then survey some of the other algorithms of this kind.

The theories considered in this chapter have, what we term, ‘effective elimination of quantifiers’. However, the qualifier ‘effective’ has very little content. In Chapter 4 we will give a more general characterization of $EQ$. The word ‘effective’ is used here merely to distinguish these examples from this later characterization.

At first sight it looks rather complicated to organize an algorithm which eliminates quantifiers from a theory. This is because we have to handle all possible combinations of quantifiers. However, some of the basic results of logic help with this organization, and takes us to the heart of the problem.
2. The effective elimination of quantifiers

2.2 THEOREM. To eliminate quantifiers for a theory $T$ it is sufficient (and necessary) to find a quantifier-free equivalent (relative to $T$) of each formula

$$(\exists w)\delta(w, v_1, \ldots, v_k)$$

where $\delta$ is a conjunction of literals (in the indicated variables) and where the quantified variable $w$ occurs in each such literal.

**Proof.** Suppose we can eliminate the quantifier $(\exists w)$ from each formula of the indicated kind. We show how to eliminate quantifiers from progressively larger classes of formulas until we have dealt with all formulas.

Consider first a formula

$$(\exists w)[\gamma \land \delta]$$

where each of $\gamma$ and $\delta$ is a conjunction of literals, where $w$ does not occur in $\gamma$, but $w$ does occur in each conjunct of $\delta$. This formula is logically equivalent to

$$\gamma \land (\exists w)\delta$$

so, by the given algorithm, we can eliminate the quantifier $(\exists w)$. In other words, we can eliminate the quantifier from any formula

$$(\exists w)\delta$$

where $\delta$ is any conjunction of literals (without any restrictions on the occurrences of $w$).

Consider any formula

$$(\exists w)[\delta_1 \lor \ldots \lor \delta_m]$$

where each $\delta_i$ is a conjunction of literals. (Thus every quantifier-free formula can be put in the disjunctive normal form of this matrix.) The whole formula is logically equivalent to

$$(\exists w)\delta_1 \lor \ldots \lor (\exists w)\delta_m$$

so, by the above algorithm, we can eliminate the quantified variable from each of these separate disjuncts. In other words, we can eliminate the quantifier from any formula

$$(\exists w)\delta$$

where $\delta$ is any quantifier-free formula.

Consider any formula

$$(\exists w_1, \ldots, w_1)\delta$$

where $\delta$ is quantifier-free. By considering

$$(\exists w_1)\delta \quad (\exists w_2, w_1)\delta \quad \cdots \quad (\exists w_t, \ldots, w_1)\delta$$

we can eliminate each quantifier in turn (from the inside) using the quantifier-free equivalents at the successive stages. In other words, we can eliminate the quantifiers from any formula

$$(\exists w)\delta$$

where $\delta$ is quantifier-free and $(\exists w)$ is any block of existentially quantified variables.
We now eliminate the quantifiers from each $\exists_{n+1}$-formula. We proceed by recursion on $n$. The base case, $n = 0$, is dealt with above. For the recursion step, $n \mapsto n + 1$, consider any $\exists_{n+2}$-formula

$$\psi = (\exists w)\phi$$

where $\phi$ is a $\forall_{n+1}$-formula and $w$ is a list of variables. The negation $\neg \phi$ is a $\exists_{n+1}$-formula so, by recursion, we obtain

$$T \vdash \neg \phi \leftrightarrow \delta$$

for some quantifier-free formula $\delta$. In particular.

$$T \vdash \phi \leftrightarrow \neg \delta$$

so that

$$T \vdash \psi \leftrightarrow (\exists w)\neg \delta$$

and it suffices to apply the base algorithm to eliminate this last block of quantifiers. $\blacksquare$

In the next two sections we look at two particular theories, and show that each has $EQ$.

Exercises

2.1 Let $T$ be a theory with $EQ$ formalized in language with a finite signature where there are no operation symbols.

What can you say about the size of the boolean algebra of sentences modulo $T$?

How many complete extensions does $T$ have?

2.2 Linear orders

In this section we use a signature with just one symbol, a 2-placed relation symbol which we write as an infix.

We look at structures

$$\mathfrak{A} = (A, \leq)$$

where $\leq$ is a special kind of linear ordering of the carrier $A$.

2.3 DEFINITION. A line is a structure $(A, \leq)$ which is a dense linear order without end points. $\blacksquare$

In other words, a line (in this sense) is a linearly ordered set with no first point, no last point, and with no gaps, that is between each pair of distinct points there is a third point. The terminology ‘line’ is not standard but it is a lot more convenient than ‘dense linear order without end points’.

There are two important examples of lines.

2.4 EXAMPLE. Both $\mathbb{Q}$ and $\mathbb{R}$ (with their natural orderings) are lines. In section 5.1 we see that $\mathbb{Q}$ is the only countable line (up to isomorphism). $\blacksquare$
In Chapter 5.1 we show that \( Q \) is the only countable line (up to isomorphism). To do that we use one of the most important techniques in model theory.

By writing down the appropriate axioms, it is easy to see that the class of lines is elementary. Furthermore, the theory \( T \) of this class is \( \forall_2 \)-axiomatizable with a finite set of axioms. We need not write down all of these axioms, but we should look at some of them.

Each line \( \mathfrak{A} \) is a poset which is linear. The distinguished attribute \( \leq \) is reflexive, and (the universal closure of)

\[
(u \leq v) \lor (v \leq u)
\]

is the axiom which ensures linearity. It can be checked that

\[
(v \nleq u) \iff (u \leq v) \land (u \neq v) \quad (u \leq v) \iff (v \nleq u) \land (u \equiv v)
\]

are consequence of this and the other axioms. It is convenient to let

\[
u < v \quad \text{abbreviate} \quad v \nleq u
\]

and pretend that this is an atomic formula. In fact, we could axiomatize the class using a signature with two 2-placed relations \( \leq \) and \( < \), and add

\[
(u < v) \iff (v \nleq u)
\]

as an axiom. In some ways that is neater. Notice that

\[
(u \neq v) \iff (u < v) \lor (v < u)
\]

is a consequence of these axioms.

So far we have used only the axioms of linearly ordered sets, and this leads to a useful observation.

### 2.5 Lemma

Relative to the theory of linearly ordered sets, each quantifier-free formula is equivalent to a \( \{ \land, \lor \} \)-combination of formulas

\[
(u \leq v) \quad (u < v) \quad (u \equiv v)
\]

for appropriate variables \( u, v \).

In other words, provided we work in terms of both \( \leq \) and \( < \), then we can get rid of all uses of negation. In fact, we can go further. Since

\[
(u \equiv v) \iff (u \leq v) \land (v \leq u)
\]

we can get rid of all uses of the equality symbol.

All these axioms are \( \forall_1 \)-sentences. However, a line has no first point and no last point, so both of

\[
(\forall w)(\exists u)[u < w] \quad (\forall w)(\exists v)[w < v]
\]

are required as further axioms. It has no gaps, so

\[
(\forall u, v)[(u < v) \rightarrow (\exists w)[(u < w) \land (w < v)]]
\]

is another axiom. These extra axioms are \( \forall_2 \)-sentences.
2.6 **THEOREM.** The theory $T$ of lines has $EQ$.

Proof. How can we eliminate the bound quantifier from a formula
\[ \theta := (\exists w)[L_1 \land \cdots \land L_i] \]
where the matrix is conjunction of literals? From the remarks above we may assume that each conjunct has one of the shapes
\[
(u \leq w) \quad (u < w) \quad (w < v) \quad (w \leq v)
\]
where $w$ is the distinguished variable (the one we want to eliminate) and $u, v$ are other variables. Of course, several other variables may occur throughout the conjuncts.

We need to consider three cases.

Suppose only conjuncts of the shapes
\[
(u \leq w) \quad (u < w)
\]
occur. Then, since a line has no last point, we have
\[ T \vdash \theta \leftrightarrow \text{true} \]
and we are done.

Suppose only conjuncts of the shapes
\[
(w < v) \quad (w \leq v)
\]
occur. Then, since a line has no first point, we have
\[ T \vdash \theta \leftrightarrow \text{true} \]
and we are done.

It remains to deal with the case where there is a mixture of ‘left’ and ‘right’ conjuncts. Consider all pairs of such conjuncts, one from the left and one from the right. There are four kinds of such pairs as listed below.

<table>
<thead>
<tr>
<th>Pair</th>
<th>Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(u \leq w) \land (w \leq v)$</td>
<td>$(u \leq v)$</td>
</tr>
<tr>
<td>$(u \leq w) \land (w &lt; v)$</td>
<td>$(u &lt; v)$</td>
</tr>
<tr>
<td>$(u &lt; w) \land (w \leq v)$</td>
<td>$(u &lt; v)$</td>
</tr>
<tr>
<td>$(u &lt; w) \land (w &lt; v)$</td>
<td>$(u &lt; v)$</td>
</tr>
</tbody>
</table>

We replace each such pair by the indicate formula, and let $\gamma$ be the conjunction of these replacements. The variable $w$ does not occur in $\gamma$. Remembering that a line has no gaps, we see that we have
\[ T \vdash \theta \leftrightarrow \gamma \]
and we are done. The implication $\theta \rightarrow \gamma$ is trivial, but the converse $\gamma \rightarrow \theta$ requires several moment’s thought.

Because of the nature of the language (in any of its variants) we quickly obtain the following.

2.7 **COROLLARY.** The theory of lines is complete.

As we will see later, linearly ordered sets and lines crop up in some unexpected places. So don’t dismiss this section as a rather routine example.
Exercises

2.2 Write down the axioms for the theory of lines and check quantifier complexity of each.

2.3 Verify the assertions between Example 2.4 and Lemma 2.5.
   Prove Lemma 2.5.

2.4 Complete the details for the end of the proof of Theorem 2.6.

2.5 Prove Corollary 2.7.

2.6 Consider the class of all structures \((A, \leq, a, b)\) where \((A, \leq)\) is a dense linearly ordered set with first point \(a\) and last point \(b\). Write down a set of axioms for this class, and show that the corresponding theory has \(EQ\).

2.7 Show that, as linearly ordered sets, \(\mathbb{Q} \equiv \mathbb{R}\).
   What does this tell you about the (Dedekind) completeness of \(\mathbb{R}\)?

   The results of the following exercises are dealt with in more detail later, but it won’t do you much harm to have a worry about them now.

2.8 Show that each linearly ordered set can be embedded in a line.

2.9 Show that the theory of lines is \(\aleph_0\)-categorical, but is not \(\kappa\)-categorical for any cardinal \(\kappa > \aleph_0\).

2.3 The natural numbers

How can we characterize the natural numbers? Dedekind observed that the structure

\[ \mathfrak{N} = (\mathbb{N}, S, 0) \]

(where \(S\) is the successor operation) is characterized by the induction property.

Each subset \(X\) of \(\mathbb{N}\) which contains 0 and is closed under \(S\) must be the whole of \(\mathbb{N}\)

Peano pointed out that some care must be taken with this idea, for we need to know which sets \(X\) are ‘acceptable’. In the present context this means that the nature of the language in which the characterization is formalized has a significant impact on the result. For instance, by Theorem 1.28, if we use a first order language, then a characterization up to isomorphism is impossible. The best we can hope for is a characterization up to elementary equivalence. In other words we can not hope for much more than a characterization of \(Th(\mathfrak{N})\).

Let’s attempt to axiomatize this theory.

There are two trivial axioms.

\[ (0) \quad (\forall v)[(Sv \neq 0)] \quad (1) \quad (\forall u, v)[(Su = Sv) \rightarrow (u = v)] \]

which are the first two of Dedekind’s axioms.
Next we want to add to these some analogue of the induction axiom (as stated above). We cannot formalize this directly in our first order language since it involves a quantification over subsets of the carrier. However, many such subsets can be named in the language, and we can certainly state the induction property for each one of these.

Let \( \phi(u_1, \ldots, u_n, v) \) be any formula in the indicated variables. We can think of this as a name for the set of all \( v \) for which the formula holds. Of course, this set depends on the parameters \( u_1, \ldots, u_n \). In other words, the formula gives us a parameterized family of subsets. We can thus regard

\[
(\forall u_1, \ldots, u_n)[\phi(u, 0) \land (\forall v)[\phi(u, v) \rightarrow \phi(u, Sv)]. \rightarrow (\forall v)\phi(u, v)]
\]

as a statement of the induction property for this particular family of subsets.

2.8 DEFINITION. Consider the language which has one constant 0 and one function symbol \( S \) where this is 1-placed.

Let \( T^+ \) be the theory in this language axiomatized by the two trivial axioms (0,1) together with all the induction axioms for all possible formulas \( \phi \).

Certainly we have \( \mathfrak{N} \models T^+ \) and we think of this as the standard model. In fact, \( \mathfrak{N} \) is a part of every model of \( T^+ \).

To analyse \( T^+ \) the first thing to do is to extract some useful consequences of the axioms. For this we need a bit of notation.

The terms of this language have a simple form. Each has one of the shapes

\[ S^k0 \quad S^kw \]

where \( w \) is a variable and \( k \in \mathbb{N} \). Here ‘\( S^k \)’ indicates a \( k \)-fold application of \( S \) to either 0 or \( w \). We abbreviate

\[ S^k0 \quad \text{by} \quad \uparrow k \]

to obtain the numerals. Thus in the structure \( \mathfrak{N} \) the numeral \( \uparrow k \) is the canonical name of \( k \in \mathbb{N} \). Notice that \( \uparrow 0 \) and 0 are the same term.

Using a mixture of internal and external induction we obtain the following.

2.9 LEMMA. We have both

\[ T^+ \vdash (\forall v)[(v \equiv \uparrow 0) \lor (\exists w)[Sw \equiv v]] \quad T^+ \vdash (\forall v)[S^{k+1}v \neq v] \]

for each \( k \in \mathbb{N} \).

With this we can produce a more amenable theory.

2.10 DEFINITION. Let \( T \) be the theory axiomatized by

\begin{align*}
(0) \quad & (\forall v)[(Sv \neq 0)] \\
(1) \quad & (\forall u,v)[(Su \equiv Sv) \rightarrow (u \equiv v)] \\
(2) \quad & (\forall v)[(v \equiv \uparrow 0) \lor (\exists w)[Sw \equiv v]]
\end{align*}
Observe that $T$ is $\forall_2$-axiomatizable. In fact, only the third axiom is a $\forall_2$-sentence, each of the others is a $\forall_1$-sentence. Lemma 2.9 shows that $T \subseteq T^+$. In particular, $\mathfrak{N} \models T$. We will show that $T$ has $EQ$ and hence, as a result, $T$ is complete, so that $T = T^+ = Th(\mathfrak{N})$.

We need some more consequences of these axioms.

2.11 Lemma. We have

$$T \vdash (\forall v)[(\exists w)[S^{k+1}w \equiv v] \leftrightarrow (v \not\equiv \lceil 0 \rceil) \wedge \cdots \wedge (v \not\equiv \lceil k \rceil)]$$

for each $k \in \mathbb{N}$.

We need one final observation before we get to the elimination algorithm. Each quantifier-free sentence of this language is equivalent to a combination of atomic sentence $(\lceil m \rceil \equiv \lceil n \rceil)$ for various $m, n \in \mathbb{N}$. Each such compound sentence is either true or false in $\mathfrak{N}$. In fact, for each such sentence $\sigma$, either $\sigma \in T$ or $\lnot \sigma \in T$.

With this we can show how to eliminate the quantifiers relative to the theory $T$.

2.12 Theorem. The theory $T$ has $EQ$.

Proof. How can we eliminate the bound variable $w$ from the formula

$$\theta := (\exists w)[L_1 \land \cdots \land L_l]$$

where each conjunct $L$ is a literal? On general ground we may assume that $w$ occurs in each $L$. Thus each such literal is either an atomic formula or the negation of an atomic formula of the shape

$$(S^m w \equiv S^n t)$$

where $m, n \in \mathbb{M}$ and the term $t$ is $w, 0$, or another variable. We consider all the various possibilities, and act accordingly.

Suppose there is an atomic formula $\alpha$, perhaps negated, of the shape

$$(S^m w \equiv S^n w)$$

for $m, n \in \mathbb{N}$. By considering the cases

$$m = n \quad m \neq n$$

we see that

$$T \vdash \alpha \leftrightarrow \text{true} \quad T \vdash \alpha \leftrightarrow \text{false}$$

holds, respectively. In other words, either that conjunct can be disregarded, or

$$T \vdash \lnot \theta$$
holds

The upshot of this is that we may assume that each occurring atomic formula has the shape

\[(S^m w \equiv s)\]

where \(m \in \mathbb{N}\) and \(w\) does not appear in the term \(s\).

Suppose one of the conjuncts \(L\) is positive, that is

\[L := (S^m w \equiv s)\]

for some term \(s\). To eliminate the quantified variable \((\exists w)\) from \(\theta\) we combine this particular conjunct \(L\) with each other conjunct \(M\) in turn. Each such conjunct \(M\) has one of the shapes

\[(S^p w \equiv t) \quad (S^p w \not\equiv t)\]

depending on its parity. Remember that \(w\) does not appear in the term \(t\). Consider the first shape. Then, working in \(T\) we have

\[T \vdash L \land M \leftrightarrow (S^{m+p} w \equiv S^p s) \land (S^{p+m} w \equiv S^m t)\]

\[\leftrightarrow (S^{m+p} w \equiv S^p s) \land (S^p s \equiv S^m t) \leftrightarrow L \land (S^p s \equiv S^m t)\]

where now \(w\) does not appear in the second component. The case where \(M\) is negative can be handled in the same way, so we obtain

\[T \vdash L \land M \leftrightarrow L \land (S^p s \equiv S^m t) \quad T \vdash L \land M \leftrightarrow L \land (S^p s \not\equiv S^m t)\]

for the positive case and negative case, respectively. From this we see that the matrix of \(\theta\) is equivalent to

\[L \land \delta\]

for some quantifier-free formula \(\delta\) in which \(w\) does not occur. Thus

\[T \vdash \theta \leftrightarrow ((\exists w)L) \land \delta\]

and it is now easier to eliminate this quantified variable.

Look at the shape of \(L\). We have

\[s = S^m r\]

where \(r\) is 0 or another variable, and \(n \in \mathbb{N}\). We need to consider whether \(n \geq m\) or \(n < m\). Setting

\[n = m + k \quad m = n + k + 1\]

as appropriate, we have

\[T \vdash L \leftrightarrow (w \equiv S^k r) \quad T \vdash L \leftrightarrow (S^{k+1} w \equiv r)\]

respectively. But

\[T \vdash (\exists w)[w \equiv S^k r] \leftrightarrow \text{true} \quad T \vdash (\exists w)[S^{k+1} w \equiv r] \leftrightarrow (r \not\equiv \{0\}) \land \cdots \land (r \not\equiv \{3\})\]
where the second equivalence come from Lemma 2.11. In either case we see that $(\exists w)L$ is equivalent to a quantifier-free formula, and hence $\theta$ is equivalent to a quantifier-free formula.

This procedure works if there is at least one positive conjunct $L$. It remains to deal with the case where each conjunct is negative. In this case $\neg \theta$ is equivalent to

$$(\forall w)[M_1 \lor \cdots \lor M_l]$$

where each disjunct $M$ has the shape

$$(S^m w \equiv S^n t)$$

where $m, n \in \mathbb{N}$ and the term $t$ is 0 or a different variable.

Since $\neg \theta$ is universally quantified we may instantiate $w$ by any numeral we please to obtain

$$T \vdash \neg \theta \rightarrow \beta$$

where $\beta$ is a quantifier-free formula in which $w$ does not occur. We may do this for a selection of numerals to obtain

$$T \vdash \neg \theta \rightarrow (\beta_0 \land \beta_1 \land \cdots \land \beta_l)$$

for appropriate $\beta_0, \beta_1, \ldots, \beta_l$. (The number of selections here, $1 + l$, is deliberately chosen so that in a moment we may use a pigeon hole argument.) We show how to select the instantiating numerals so that

$$T \vdash (\beta_0 \land \beta_1 \land \cdots \land \beta_l) \rightarrow \text{false}$$

and hence

$$T \vdash \theta$$

holds.

Consider any disjunct $M$. This has the shape

$$(S^m w \equiv S^n t)$$

for some $m, n \in \mathbb{N}$. Consider any $k \geq n$. We may set $w = \ulcorner k \urcorner$ so that the instantiated disjunct is equivalent to

$$(t \equiv \ulcorner a \urcorner)$$

for some $a \in \mathbb{N}$. In fact, $m + k = n + a$. In the same way, by setting $w = \ulcorner k + 1 \urcorner$ this instantiation of the same disjunct is equivalent to

$$(t \equiv \ulcorner a + 1 \urcorner)$$

for the same $a$ as before.

By setting $w = \ulcorner k \urcorner$ for some sufficiently large $k$ we obtain

$$T \vdash \neg \theta \rightarrow (t_1 \equiv \ulcorner a_1 \urcorner) \lor \cdots \lor (t_l \equiv \ulcorner a_l \urcorner)$$

for some $a_1, \ldots, a_l \in \mathbb{N}$. 

2. The effective elimination of quantifiers
By setting $w = \mathbf{k} + \Gamma$ we obtain

$$T \vdash \neg \theta \rightarrow (t_1 \equiv \mathbf{r}a_1 + \Gamma) \lor \cdots \lor (t_l \equiv \mathbf{r}a_l + \Gamma)$$

for the same $a_1, \ldots, a_l \in \mathbb{N}$.

Repeating this for each of

$$w := \mathbf{k}, w := \mathbf{k} + \Gamma, \ldots, w := \mathbf{k} + \Gamma,$$

we obtain

$$T \vdash \neg \theta \rightarrow \gamma$$

where $\gamma$ is

$$(t_1 \equiv \mathbf{r}a_1 + 0) \lor \cdots \lor (t_1 \equiv \mathbf{r}a_l + 0)$$

$$\land$$

$$(t_1 \equiv \mathbf{r}a_1 + 1) \lor \cdots \lor (t_1 \equiv \mathbf{r}a_l + 1)$$

$$\land$$

$$\vdots$$

$$\land$$

$$(t_1 \equiv \mathbf{r}a_1 + \Gamma) \lor \cdots \lor (t_1 \equiv \mathbf{r}a_l + \Gamma)$$

for some $a_1, \ldots, a_l \in \mathbb{N}$. This formula is a conjunction of disjunctions. We may rephrase it as a disjunction of conjunctions. Each such conjunction has the shape

$$\delta := (t_{j(0)} \equiv \mathbf{r}a_{j(0)} + 0) \land (t_{j(1)} \equiv \mathbf{r}a_{j(1)} + 1) \land \cdots \land (t_{j(l)} \equiv \mathbf{r}a_{j(l)} + \Gamma)$$

where the indexes $j(0), j(1), \ldots, j(l)$ are selected from $\{1, \ldots, l\}$. Each conjunction arises from a different selection of indexes.

For each such $\delta$ there are $1 + l$ indexes $j(\cdot)$ selected from a set of size $l$. Thus $j(r) = j(s) = j$ (say) for some $r \neq s$, and hence

$$T \vdash \delta \rightarrow (\mathbf{r}a_j + r \equiv \mathbf{r}a_j + s) \rightarrow (r \equiv s) \rightarrow \text{false}$$

which leads to

$$T \vdash \gamma \rightarrow \text{false}$$

as required.

An important by-product of this result is that we now have a complete axiomatization of the theory $Th(\mathfrak{N})$, and, what is more, we have got rid of the induction axioms. This axiomatization enables us to give a full description of all the structures $\mathfrak{A} \equiv \mathfrak{N}$. Which is nice.

We will return to this theory in Chapter 5 where we will produce a nicer version.

**Exercises**

2.10 Prove Lemma 2.9.

2.11 Prove Lemma 2.11.
2.12 (a) Show that each model of $T$ consists of a single copy of $\mathfrak{N}$ together with a family of disjoint copies of $(\mathbb{Z}, S, 0)$.
(b) Show that $T$ is $\kappa$-categorical for each uncountable $\kappa$.
(c) Describe the spectrum of countable models of $T$.

2.13 (a) Show that
\[ \mathfrak{N} \models \delta \iff T \vdash \delta \]
holds for each quantifier-free sentence $\delta$.
(b) Show that $T = Th(\mathfrak{N})$.

2.14 Exercise 2.13 shows that Definition 2.10 provides a simple axiomatization of $Th(\mathfrak{N})$. But Gödel’s incompleteness theorem says there is no such axiomatization. Explain this.

2.4 Some other examples – To be done

There are several other theories of structures based on $\mathbb{N}$ or $\mathbb{Z}$ which have $EQ$. A discussion of these is given in pages 307 – 334 of [?]

Locate, or work out, the algorithm for the theory of two successors

Chapter 4 of [?] is devoted to $EQ$. It discusses several other theories of linear orderings with $EQ$, as well as some more sophisticated theories.

Pages 49 – 60 of [?] considers theories with $EQ$. The exercises give a a fairly long list of examples (without solutions).

It would be nice to have a comprehensive list of examples. I don’t know of such a list, but presumably there is one somewhere.

Exercises

[Some needed]
3
Basic methods

We are now in a position to begin the development of model theory proper. This is the use of the compactness property to obtain information about structures and classes of structures.

3.1 Some semantic relations

We start with a description of the fundamental notions. We have met some of these earlier, but it is as well to have them defined all in one place.

3.1 Definition. Let $\mathfrak{A}, \mathfrak{B}$ be structures (for the same language).

We write

$$\mathfrak{A} \equiv \mathfrak{B}$$

and say $\mathfrak{A}$ and $\mathfrak{B}$ are elementarily equivalent if

$$\mathfrak{A} \models \sigma \iff \mathfrak{B} \models \sigma$$

holds for each sentence $\sigma$ (of the underlying language).

For each $n < \omega$ we write

$$\mathfrak{A} \equiv_n \mathfrak{B}$$

and say $\mathfrak{A}$ and $\mathfrak{B}$ are $n$-equivalent if

$$\mathfrak{A} \models \sigma \iff \mathfrak{B} \models \sigma$$

holds for each $\forall_n$-sentence $\sigma$ (of the underlying language).

It is easy to check that each of these relations is an equivalence relation. We often write

$$\equiv_\omega \text{ for } \equiv \quad \equiv_\infty \text{ for } \cong$$

and this allows us to treat certain properties of a whole family of equivalence relations in one go.

Notice that if for each sentence we have

$$\mathfrak{A} \models \sigma \implies \mathfrak{B} \models \sigma$$

then $\mathfrak{A} \equiv \mathfrak{B}$. Similarly if this implication holds for each quantifier-free sentence, then $\mathfrak{A} \equiv_0 \mathfrak{B}$. However, a similar observation fails for $\equiv_n$ for $n \neq 0$. 

q q
3.2 DEFINITION. Let \( \mathfrak{A}, \mathfrak{B} \) be structures (for the same language). For each \( n < \omega \) we write
\[
\mathfrak{A} \equiv (\forall_n) \mathfrak{B} \quad \mathfrak{A} \equiv (\exists_n) \mathfrak{B}
\]
if
\[
\mathfrak{A} \models \sigma \iff \mathfrak{B} \models \sigma
\]
for each \( \forall_n \)-sentence \( \exists_n \)-sentence \( \sigma \), respectively.

Thus \( \equiv (\forall_0), \equiv (\exists_0) \) and \( \equiv_0 \) are the same relation, but \( \equiv (\forall_{n+1}) \) and \( \equiv (\exists_{n+1}) \) are converse relations.

Recall that, from Definition 1.10, we have the notion
\[
\mathfrak{A} \subseteq \mathfrak{B}
\]
of one structure being a substructure of another (or one structure being a superstructure of another). We can refine this notion.

3.3 DEFINITION. Let \( \mathfrak{A} \subseteq \mathfrak{B} \) be structures (for the same language).

We write
\[
\mathfrak{A} \prec \mathfrak{B}
\]
and say \( \mathfrak{A} \) is an elementary substructure of \( \mathfrak{B} \) if
\[
\mathfrak{A} \models \phi(a) \iff \mathfrak{B} \models \phi(a)
\]
for each formula \( \phi(v) \) (of the underlying language) and point \( a \) of \( \mathfrak{A} \) which matches the batch \( v \) of free variables of \( \phi \).

For each \( n < \omega \) we write
\[
\mathfrak{A} \prec_n \mathfrak{B}
\]
and say \( \mathfrak{A} \) is a \( n \)-substructure of \( \mathfrak{B} \) if
\[
\mathfrak{A} \models \phi(a) \iff \mathfrak{B} \models \phi(a)
\]
for each \( \forall_n \)-formula \( \phi(v) \) (of the underlying language) and point \( a \) of \( \mathfrak{A} \) which matches batch \( v \) of free variables of \( \phi \).

Thus \( \prec_0 \) and \( \subseteq \) and are the same relation. As with \( \equiv \) we sometimes write
\[
\prec_\omega \quad \text{for} \quad \prec
\]
so that we can treat certain properties of the relations together. (We don’t have a relation \( \prec_\infty \) since \( \mathfrak{A} \prec_\infty \mathfrak{B} \) could only mean \( \mathfrak{A} = \mathfrak{B} \).

In Section 1.2 we mentioned the informal notion of an isomorphism between two structure. We can now make that precise, and set up the more general notion of an embedding of one structure into another. As often happens this is rather tedious in the initial stages, but it becomes simpler later on.
3.4 DEFINITION. Let $\mathfrak{A}, \mathfrak{B}$ be structures (for the same language).

An embedding

$$\begin{array}{c}
\mathfrak{A} \\
\frown \\
\mathfrak{B}
\end{array}$$

(from $\mathfrak{A}$ to $\mathfrak{B}$) is a function

$$f : A \longrightarrow B$$

between the carriers, as indicated, which is injective and with the following properties.

- For each constant symbol $K$ of the signature

  $$fa = b$$

  where

  $$a = \mathfrak{A}[K] \quad b = \mathfrak{B}[K]$$

  are the input and output of $f$.

- For each $n$-placed relation symbol $R$ of the signature

  $$\mathfrak{A}[R]a_1 \cdots a_n \Longleftrightarrow \mathfrak{B}[R]b_1 \cdots b_n$$

  for each $a_1, \ldots, a_n$ from $\mathfrak{A}$ and $b_i = fa_i$ for each $1 \leq i \leq n$.

- For each $n$-placed operation symbol $O$ of the signature

  $$f(\mathfrak{A}[O]a_1 \cdots a_n) = \mathfrak{B}[O]b_1 \cdots b_n$$

  for each $a_1, \ldots, a_n$ from $\mathfrak{A}$ and $b_i = fa_i$ for each $1 \leq i \leq n$.

There is one clause for each symbol of the signature.

This looks a bit complicated, but it is saying nothing more than the range $f[A]$ of $f$ is (the carrier of a) substructure of $\mathfrak{B}$ and $f$ is an isomorphism from $\mathfrak{A}$ to this substructure. In particular, an isomorphism is a surjective embedding.

Exercises

3.1 Show that for each $r \in \mathbb{N} \cup \{\omega, \infty\}$, the relation $\equiv_r$ is an equivalence.

Show that for each $r, s \in \mathbb{N} \cup \{\omega, \infty\}$,

$$\mathfrak{A} \equiv_s \mathfrak{B} \implies \mathfrak{A} \equiv_r \mathfrak{B}$$

for $r \leq s$ (using the obvious comparison).

3.2 Show that

$$\mathfrak{A} \prec_r \mathfrak{B} \implies \mathfrak{A} \equiv (\exists_{r+1}) \mathfrak{B}$$

(for each $r < \omega$).
3.2 The diagram technique

We concluded Section 3.1 with the rather tedious definition of an embedding between structures. Using quantifier-free formulas we can characterize this notion in a more convenient way. (Before you read this proof you might like to have a look at Exercise 1.17.)

3.5 LEMMA. Let $\mathfrak{A}, \mathfrak{B}$ be a pair of structures, and let $f : A \longrightarrow B$ be a function between the carriers. Then $f$ is an embedding (of $\mathfrak{A}$ to $\mathfrak{B}$) if and only if for each quantifier-free formula $\delta(v_1, \ldots, v_n)$ we have

$$\mathfrak{A} \models \delta(a_1, \ldots, a_n) \iff \mathfrak{B} \models \delta(b_1, \ldots, b_n)$$

for each elements $a_1, \ldots, a_n$ from $\mathfrak{A}$ and where $b_i = fa_i$ for each $1 \leq i \leq n$ are the corresponding elements of $\mathfrak{B}$.

Proof. Suppose first that this semantic equivalence does hold. Then, using the atomic formulas

$$\begin{align*}
\delta(v, w) & := (v \equiv w) \\
\delta(v) & := (v \equiv K) \\
\delta(v_1, \ldots, v_n) & := Rv_1 \cdots v_n \\
\delta(v_0, v_1, \ldots, v_n) & := (v_0 \equiv O v_1, \ldots, v_n)
\end{align*}$$

we see that $f$ is injective and the various required signature clauses hold.

Conversely, suppose $f$ is an embedding. By definition, we have the required equivalence for each atomic formula $\delta$, and a simple induction gives it for each quantifier-free formula. ■

Given two structures $\mathfrak{A}, \mathfrak{B}$, when can we say $\mathfrak{A}$ is embeddable in $\mathfrak{B}$? In other words, when does an embedding

$$\mathfrak{A} \longrightarrow \mathfrak{B}$$

exits? There is an important model theoretic technique – the diagram technique – which handles this and similar questions.

So far, at any one time we have been concerned with just one language, and for the most part this has been left in the background. We are now going to use two languages, one of which is an enrichment of the other. This must be done with some care, and since this is the first time we have used this trick, we will take it slowly.

Let $L$ be the parent language (so that $\mathfrak{A}$ and $\mathfrak{B}$, above, are $L$-structures). We enrich $L$ by adding a family of new constant symbols to the signature. This generates a larger language $L'$

The new constant symbols are called the parameters of the enriched language $L'$ (to distinguish them from the old constant symbols of $L$ which also occur in $L'$). Furthermore, each parameter has an intended interpretation.

Let $a$ be an enumeration of a part of, or the whole of, $\mathfrak{A}$. We enrich $L$ by adding a name for each one of the elements in $a$. These names are the parameters.

Strictly speaking we should distinguish between the element $a$ of $\mathfrak{A}$ and the parameter, new constant, introduced to name $a$. We should write something like $K_a$ for this
3.2. The diagram technique

parameter. However, in practice, we do not distinguish between an element \( a \) in \( \mathfrak{A} \) and its name which is added to \( L \). In particular, we write \( L(a) \) for the enriched language. Furthermore, when \( a \) enumerates the whole of \( \mathfrak{A} \), as it often does, we write \( L(\mathfrak{A}) \) for the enriched language.

Of course, this rather sloppy convention has some obvious pit-falls. However these are easily avoided once we become familiar with the convention and know how it is used. (The alternative is worse, and attractive only to the analytically retentive.)

The idea is that \( L(a) \) is designed to talk about \( \mathfrak{A} \), and specifically about the part enumerated by \( a \). Thus the structure \((\mathfrak{A}, a)\) (which is formed from \( \mathfrak{A} \) by distinguishing each element in \( a \)) is a kind of canonical structure for \( L(a) \). Of course, this language can talk about other structures of the form \((\mathfrak{B}, b)\) where \( \mathfrak{B} \) is an \( L \)-structure and \( b \) is an enumeration of certain elements of \( \mathfrak{B} \).

3.6 DEFINITION. Let \( \mathfrak{A} \) be an \( L \)-structure, let \( a \) be an enumeration of a part of \( \mathfrak{A} \), and consider the enriched language \( L(a) \).

The a-diagram

\[
\text{Diag}(\mathfrak{A}, a)
\]

of \( \mathfrak{A} \) is the set of quantifier-free \( L(a) \)-sentences which hold in \((\mathfrak{A}, a)\).

When \( a \) enumerates the whole of \( \mathfrak{A} \) we call \( \text{Diag}(\mathfrak{A}, a) \) the diagram of \( \mathfrak{A} \).

What does a sentence in \( \text{Diag}(\mathfrak{A}, a) \) look like? Each \( L(a) \)-sentence has the shape

\[
\delta(a_1, \ldots, a_n)
\]

where

\[
\delta(v_1, \ldots, v_n)
\]

is a quantifier-free \( L \)-formula in the indicated variable, and \( a_1, \ldots, a_n \) are taken from \( a \). If we collapse these two finite list to a batch \( v \) and point \( a \), then each member of \( \text{Diag}(\mathfrak{A}, a) \) has the shape \( \delta(a) \). The crucial condition is that

\[
\mathfrak{A} \models \delta(a)
\]

and we select all such \( L(a) \)-sentences.

In this way we can flit between the two languages \( L \) and \( L(a) \), and creatively confuse syntax with semantics. The following result illustrates how this is used. If at first you find the proof confusing then be pernickety and distinguish between an element \( a \) of \( \mathfrak{A} \) and the parameter \( K_a \) chosen to name \( a \).

3.7 LEMMA. Let \( \mathfrak{A}, \mathfrak{B} \) be two \( L \)-structures (for some language \( L \)). Let \( a \) be an enumeration of the whole of \( \mathfrak{A} \). Then \( \mathfrak{A} \) is embeddable in \( \mathfrak{B} \) if and only if

\[
(\mathfrak{B}, b) \models \text{Diag}(\mathfrak{A}, a)
\]

for some enumeration \( b \) of a part of \( \mathfrak{B} \).

Proof. Suppose first that \( \mathfrak{A} \) is embeddable in \( \mathfrak{B} \). Thus we have a function of a certain kind

\[
f : A \rightarrow B
\]
from the carrier of $\mathfrak{A}$ to the carrier of $\mathfrak{B}$. By transferring each element $a$ in $a$ to the element $b = f(a)$ of $\mathfrak{B}$ we obtain an enumeration $b$ of a part of $\mathfrak{B}$. A use of Lemma 3.5 in one direction (the ‘only if’ direction) gives

$$(\mathfrak{B}, b) \models \text{Diag}(\mathfrak{A}, a)$$

as required.

Conversely, suppose $(\mathfrak{B}, b)$ does model the $a$-diagram of $\mathfrak{A}$. By sending each element $a$ in $a$ to the corresponding element $b$ in $b$ we obtain a function $f$ from $A$ to $B$. A use of Lemma 3.5 in the other direction (the ‘if’ direction) shows that $f$ is an embedding. ■

Once the above ideas have been absorbed, this result is almost a triviality. The enumeration $b$ gives the range of the required embedding, and the matching of $a$ with $b$ gives the required function.

Let’s work around this idea to make sure we understand it.

3.8 LEMMA. Let $\mathfrak{A}$ and $\mathfrak{B}$ be a pair of structures for the same language, and suppose $\mathfrak{A}$ is embeddable in $\mathfrak{B}$. Then

$$\mathfrak{B} \models \alpha \implies \mathfrak{A} \models \alpha$$

for each $\forall_1$-sentence $\alpha$ (of the underlying language).

Proof. Let $f$ be the given embedding of $\mathfrak{A}$ into $\mathfrak{B}$. Suppose

$$\mathfrak{B} \models \alpha$$

for some $\forall_1$-sentence $\alpha$. This sentence has the shape (is equivalent to)

$$(\forall v)\delta(v)$$

where $v$ is a batch of variables and $\delta(v)$ is a quantifier-free formula. Consider any point $a$ of $\mathfrak{A}$ which matches the batch $v$. We must show that

$$\mathfrak{A} \models \delta(a)$$

holds.

Using the embedding $f$ we may transfer $a$ to $\mathfrak{B}$ to obtain a point $b$ of $\mathfrak{B}$. This point $b$ matches $v$. Since $\mathfrak{B} \models \alpha$ we have

$$\mathfrak{B} \models \delta(b)$$

and then a use of Lemma 3.5 gives the required result. ■

Lemmas 3.5, 3.7, and 3.8 are concerned with the comparison of two given structures. A different use of the same ideas shows the existence of a structure relative to a given one.

3.9 LEMMA. For a language $L$ let $\Sigma$ be a set of sentences of $L$, let $\mathfrak{A}$ be a structure for $L$, and suppose

$$\Sigma \vdash \alpha \implies \mathfrak{A} \models \alpha$$

for each $\forall_1$-sentences $\alpha$. Then $\mathfrak{A} \subseteq \mathfrak{B}$ for some model $\mathfrak{B}$ of $\Sigma$. 

Proof. Let $a$ be an enumeration of $A$ and consider the diagram $\text{Diag}(A, a)$. We show first that 
\[ \Sigma \cup \text{Diag}(A, a) \]
is consistent.

By way of contradiction suppose that this set is not consistent. By an application of compatness, Theorem 1.25 or Theorem 1.27, some finite part of this set in not consistent. Thus there is a quantifier-free $L$-formula $\delta(v)$ and a point $a$ of $A$ matching $v$ such that both 
\[ \Sigma \cup \{\delta(a)\} \text{ is not consistent} \quad A \models \delta(a) \]
hold. We now flit between the two views of $a$. On the left as a list of constants in the enriched language $L(a)$, and on the right as a point of $A$.

The left hand condition ensures that 
\[ \Sigma \vdash \neg \delta(a) \]
holds. Here $\vdash$ is the consequence relation for the language $L(a)$. The crucial property is that the constants $a$ do not occur in $\Sigma$ and so can be interpreted in any way whatsoever without any restricting demands from $\Sigma$. In other words, for this particular relationship, the constants $a$ behave exactly like free variables, so we may replace them by free variable of the language $L$ to show that 
\[ \Sigma \vdash \neg \delta(v) \]
holds.

Since $\Sigma$ is a set of sentences this gives 
\[ \Sigma \vdash (\forall v) \neg \delta(v) \]
and hence 
\[ A \models (\forall v) \neg \delta(v) \]
by the given relationship between $\Sigma$ and $A$.

We now take the other view of $a$, that of a point of $A$. Since $a$ matches $v$ we can instantiate to get 
\[ A \models \neg \delta(a) \]
which gives us the required contradiction.

We now know that $\Sigma \cup \text{Diag}(A, a)$ is consistent and hence Lemma 3.7 gives a model $B$ of $\Sigma$ together with and embedding 
\[ A \longrightarrow B \]
from $A$.

This is not quite what we want. We require the embedding to be an insertion. To achieve that we observe that we may replace $B$ by any structure isomorphic to $B$. This new structure will still be a model of $\Sigma$, and the original embedding followed by the isomorphism is still an embedding. Thus all we need to do is replace the image of $A$ in $B$ by $A$ itself.

We made rather heavy weather of this proof by describing almost every detail. We did this to show exactly how it works. It uses the essential trick of the method of diagrams. In future similar proofs will not be so detailed. At first this may cause you some
consternation. If so, simply write out the proof in the detailed finicky manner. You will soon get used to the idea.

We now introduce a notion that is used throughout model theory. With this notion we can reformulate Lemma 3.9 in a more conventional style.

3.10 DEFINITION. Let $T$ be a $L$-theory, and let $\mathfrak{A}$ be a $L$-structure (for some language $L$). We set

$$T[\mathfrak{A}] = T \cup \text{Diag}(\mathfrak{A}, a)$$

where $a$ is an enumeration of the whole of $\mathfrak{A}$ to produce a set of $L(\mathfrak{A})$-sentences.

Strictly speaking, $T[\mathfrak{A}]$ depends on which particular enumeration $a$ is used. However, this will never be a cause for concern. Notice that although $T$ is an $L$-theory the set $T[\mathfrak{A}]$ need not be an $L(\mathfrak{A})$-theory, and it need not be consistent.

3.11 DEFINITION. Let $T$ be a $L$-theory, and let $\mathfrak{A}$ be a $L$-structure (for some language $L$).

- (m) We say $\mathfrak{A}$ is a model of $T$ if $\mathfrak{A} \models T$. Let $\mathcal{Md}(T)$ be the class of models of $T$.
- (s) We say $\mathfrak{A}$ is a submodel of $T$ if $\mathfrak{A} \subseteq \mathfrak{B}$ for some $\mathfrak{B} \models T$. Let $\mathcal{S}(T)$ be the class of submodels of $T$.

By construction we have $\mathcal{Md}(T) \subseteq \mathcal{S}(T)$. Later we will look at various other subclasses of $\mathcal{S}(T)$.

The following is essentially the same result as Lemma 3.9.

3.12 LEMMA. Let $T$ be a $L$-theory, and let $\mathfrak{A}$ be a $L$-structure (for some language $L$). The set $T[\mathfrak{A}]$ of $L(\mathfrak{A})$-sentences is consistent if and only if $\mathfrak{A} \in \mathcal{S}(T)$.

Proof. If $\mathfrak{A} \in \mathcal{S}(T)$ then $\mathfrak{A} \subseteq \mathfrak{B}$ for some $\mathfrak{B} \models T$, and this provides the required model of $T[\mathfrak{A}]$.

Conversely, suppose $T[\mathfrak{A}]$ is consistent, and let $(\mathfrak{B}, b)$ be any model. Then $(\mathfrak{B}, b) \models T$, and hence $\mathfrak{B} \models T$ (since the parameters don’t appear in $T$). Since $(\mathfrak{B}, b) \models \text{Diag}(\mathfrak{A}, a)$, there is an embedding $\mathfrak{A} \longrightarrow \mathfrak{B}$. By replacing $\mathfrak{B}$ by a suitable isomorphic copy, we may suppose $\mathfrak{A} \subseteq \mathfrak{B}$, and hence $\mathfrak{A} \in \mathcal{S}(T)$.

Notice how short this proof is. This is the way a diagram proof is normally written out. We hide all the flitting about between parameter and variable. You should try to get use this style. Otherwise you might be here until way after closing time.

So far we have considered only embeddings and the quantifier-free diagram of a structure. There are various refinements which are sometimes useful.

3.13 DEFINITION. Let $\mathfrak{A}, \mathfrak{B}$ be a pair of structures, and let

$$f : A \longrightarrow B$$

be a function between the carriers. Then, for each $r < \omega$ or $r = \omega$, we say $f$ is a $\prec_r$-embedding if and only

$$\mathfrak{A} \models \phi(a_1, \ldots, a_n) \iff \mathfrak{B} \models \phi(b_1, \ldots, b_n)$$

for each $\forall_y$-formula $\phi(v_1, \ldots, v_n)$ and elements $a_1, \ldots, a_n$ from $\mathfrak{A}$ and where $b_i = fa_i$ for each $1 \leq i \leq n$. 

Lemma 3.5 shows that a $\prec_0$-embedding is nothing more than an embedding. As $r$ increases the notion of a $\prec_r$-embedding becomes more and more restrictive, until the case $r = \omega$ gives the notion of an elementary embedding.

How might we produce these refined embeddings? Have a think about Exercise 3.4.

**Exercises**

3.3 This exercise generalizes Lemma 3.9.
(a) Let $\Sigma$ be a set of sentences, let $\mathfrak{A}$ be a structure, and suppose
\[ \Sigma \vdash \alpha \implies \mathfrak{A} \models \alpha \]
for each $\forall_2$-sentences $\alpha$. Show that $\mathfrak{A} \prec_1 \mathfrak{B}$ for some model $\mathfrak{B}$ of $\Sigma$.
(b) Can you generalize this to deal with $\forall_{n+1}$-sentences?

3.4 Let $\mathfrak{A}, \mathfrak{B}$ be arbitrary structures (for the same language). Show that for each $n < \omega$,
\[ \mathfrak{A} \equiv (\forall_{n+1}) \mathfrak{B} \]
holds if and only if
\[ \mathfrak{A} \xrightarrow{f} \mathfrak{C} \quad \mathfrak{B} \xrightarrow{g} \mathfrak{C} \]
for some $\prec$-embedding $f$ and some $\prec_n$-embedding $g$ to a common structure $\mathfrak{C}$.

3.5 Let $\mathfrak{A}$ and $\mathfrak{B}$ be a pair of structures suitable for the same language. Show that $\mathfrak{A} \equiv \mathfrak{B}$ if and only if
\[ \mathfrak{A} \prec \mathfrak{A}' \quad \mathfrak{B} \prec \mathfrak{B}' \]
for some pair $\mathfrak{A}', \mathfrak{B}'$ of isomorphic structures.

3.3 **Restricted axiomatization**

A set of axioms for a theory $T$ is a subset $\Sigma \subseteq T$ such that
\[ \sigma \in T \iff \Sigma \vdash \sigma \]
for each sentence $\sigma$. Every theory $T$ has a set of axioms, for we can always take $\Sigma = T$. However, we usually want a more interesting set of axioms.

3.14 **DEFINITION.** Let $T$ be a theory (in some language $L$). For each $n < \omega$, we say $T$ is $\forall_n$-axiomatizable if $T$ is the set of all consequences of some set of $\forall_n$-sentences (of $L$).

We are most often concerned with $\forall_1$-axiomatizability and $\forall_2$-axiomatizability, but the general notion is useful. Sometimes we can tell that a theory is $\forall_n$-axiomatizable by simple observing the axioms we have written down for it. On other occasions this isn’t helpful, either because we have a set of axioms that is too complicated, or we do not even have a set of axioms.
3.15 EXAMPLE. Most theories that arise in algebra are $\forall_2$-axiomatizable or better. For instance, by choosing the right signature the theory of rings is $\forall_1$-axiomatizable and the theory of fields is $\forall_2$-axiomatizable. However, is it possible that within this signature, the theory of fields has a set of $\forall_1$-axioms?

Most of the axioms of Peano arithmetic are $\forall_2$ (and even $\forall_1$). However, the usual induction axioms uses sentences of arbitrary quantifier complexity. Is there some way by which we can replace these axioms by ones of bounded quantifier complexity? ■

How can we test a theory $T$ for $\forall_n$-axiomatizability? Observe that if a theory $T$ is $\forall_n$-axiomatizable, then $T \cap \forall_n$ is such a set of axioms. Thus we need to investigate how $Md(T)$ and $Md(T \cap \forall_n)$ are related. Here is the answer for $n = 1$.

3.16 LEMMA. Let $T$ be a theory (in some language $L$). For each structure $\mathfrak{A}$ (suitable for $L$), the following

(i) $\mathfrak{A} \in S(T)$  
(ii) $\mathfrak{A} \models T \cap \forall_1$

are equivalent.

Proof. The implication $(i) \Rightarrow (ii)$ is an immediate consequence of Lemma 3.8.

The implication $(ii) \Rightarrow (i)$ is an immediate consequence of Lemma 3.9. ■

This result gives us a characterization of $\forall_1$-axiomatizability.

3.17 THEOREM. A theory $T$ (in some language $L$) is $\forall_1$-axiomatizable if any only if each submodel of $T$ is a model.

Proof. Suppose first that $T$ is $\forall_1$-axiomatizable. Then, in fact, $T \cap \forall_1$ is a set of axioms for $T$, and hence

$\mathfrak{A} \models T \cap \forall_1 \implies \mathfrak{A} \models T$

for all structures $\mathfrak{A}$. In particular, we have

$\mathfrak{A} \in S(T) \implies \mathfrak{A} \models T \cap \forall_1 \implies \mathfrak{A} \models T$

to show that $S(T) = Md(T)$.

Conversely, suppose $S(T) = Md(T)$. Then, by Lemma 3.16, we have

$\mathfrak{A} \models T \cap \forall_1 \implies \mathfrak{A} \in S(T) = Md(T) \implies \mathfrak{A} \models T$

(for each structure $\mathfrak{A}$) to show that $T \cap \forall_1$ is set of axioms of $T$. ■

The same kind of argument gives us an interpolation result.

3.18 THEOREM. For each theory $T$ and sentences $\lambda, \rho$ (in the same language), the following are equivalent.

(i) We have

$T \models \lambda \rightarrow \sigma \quad T \models \sigma \rightarrow \rho$

for some $\forall_1$-sentence $\sigma$. 
3.3. Restricted axiomatization

(ii) We have

\[ \mathcal{B} \models \lambda \Rightarrow \mathcal{A} \models \rho \]

for all models \( \mathcal{A}, \mathcal{B} \) of \( T \) with \( \mathcal{A} \subseteq \mathcal{B} \).

**Proof.** Only the implication \((ii) \Rightarrow (i)\) offers much resistance. To prove this let \( T' \) be the deductive closure of \( T \cup \{ \lambda \} \) and let \( \Sigma = T' \cap \forall_1 \). Thus, \( \Sigma \) is the set of all \( \forall_1 \)-sentence \( \sigma \) such that

\[ T \vdash \lambda \rightarrow \sigma \]

and it suffices to show that

\[ T \cup \Sigma \vdash \rho \]

holds.

To this end, consider any model \( \mathcal{A} \) of \( T \cup \Sigma \). By Lemma 3.16 there is some model \( \mathcal{B} \models T' \) with \( \mathcal{A} \subseteq \mathcal{B} \). Both \( \mathcal{A} \) and \( \mathcal{B} \) are models of \( T \) and \( \mathcal{B} \models \lambda \) (since \( \lambda \in T' \)). Thus, by (ii), we have \( \mathcal{A} \models \rho \), as required. \( \blacksquare \)

These last few results are concerned with \( \forall_1 \)-sentences. All of them can be generalized to give analogues using \( \forall_{n+1} \)-sentences for \( n < \omega \). These results are dealt with in the exercises.

The diagram technique is quite versatile and is used many times in model theory. Here is an application that produces a different kind of result. This is a variant of the result of Exercise 3.4.

3.19 LEMMA. Let \( \mathcal{A} \subseteq \mathcal{B} \) be a pair of structures (of the same signature). The following are equivalent.

(i) \( \mathcal{A} \prec_1 \mathcal{B} \)

(ii) There is some structure \( \mathcal{C} \) with \( \mathcal{B} \subseteq \mathcal{C} \) and \( \mathcal{A} \prec \mathcal{C} \).

(iii) There is some structure \( \mathcal{C} \) with \( \mathcal{B} \subseteq \mathcal{C} \) and \( \mathcal{A} \prec_1 \mathcal{C} \).

**Proof.** The two implications \((ii) \Rightarrow (iii)\) and \((iii) \Rightarrow (i)\) are immediate. The implication \((i) \Rightarrow (ii)\) is the content of this result. To prove this consider any pair \( \mathcal{A} \prec_1 \mathcal{B} \). Let \( \mathcal{a} \) be an enumeration of \( \mathcal{A} \) and let \( \mathcal{b} \) be an enumeration of \( \mathcal{B} \). (The elements of \( \mathcal{A} \) will be enumerated twice, once in \( \mathcal{a} \) and once in \( \mathcal{b} \), but this doesn’t matter.) It suffices to show that

\[ Th(\mathcal{A}, \mathcal{a}) \cup Diag(\mathcal{B}, \mathcal{a}, \mathcal{b}) \]

is consistent.

(Notice that this involves two enrichments of the underlying language; a first one by the addition of \( \mathcal{a} \), and then a further one by the addition of \( \mathcal{b} \).)

If this set is not consistent then

\[ Th(\mathcal{A}, \mathcal{a}) \vdash \neg \delta(\mathcal{b}, \mathcal{a}) \]

for some quantifier-free formula \( \delta(w,v) \) (of the underlying language) some point \( a \) of \( \mathcal{A} \) and some point \( b \) of \( \mathcal{B} \). Since \( b \) does not occur in (the language of) \( Th(\mathcal{A}, \mathcal{a}) \), we have

\[ Th(\mathcal{A}, \mathcal{a}) \vdash \neg \delta(w,a) \]
and hence

\[ \text{Th}(\mathfrak{A}, a) \vdash (\forall w) \neg \delta(w, a) \]

so that

\[ \mathfrak{A} \models (\forall w) \neg \delta(w, a) \]

holds. But now (i) gives \( \mathfrak{B} \models (\forall w) \neg \delta(w, a) \), which leads to a contradiction. \[ \blacksquare \]

This result can be generalized to produce a characterization of the relation \( \prec_{n+1} \) between structures.

**Exercises**

3.6 Let \( T \) be a theory (in some language \( L \)). Let \( n < \omega \).

(a) Show that a structure \( \mathfrak{A} \) is a model of \( T \cap \forall_{n+1} \) if and only if there is a model \( \mathfrak{B} \models T \) with \( \mathfrak{A} \prec n \mathfrak{B} \).

(b) Show that \( T \) is \( \forall_{n+1} \)-axiomatizable if and only if

\[ \mathfrak{A} \prec_{n} \mathfrak{B} \models T \implies \mathfrak{A} \models T \]

holds (for all structures \( \mathfrak{A}, \mathfrak{B} \)).

(c) Show that for each sentence \( \lambda, \mu \) the following are equivalent.

(i) There is an \( \forall_{n+1} \)-sentence \( \sigma \) such that both \( T \vdash \lambda \rightarrow \sigma \) and \( T \vdash \sigma \rightarrow \rho \) hold.

(ii) The implication

\[ \mathfrak{B} \models \lambda \implies \mathfrak{A} \models \rho \]

holds for all models \( \mathfrak{A}, \mathfrak{B} \) of \( T \) with \( \mathfrak{A} \prec_{n} \mathfrak{B} \).

3.7 Let \( \mathfrak{A} \subseteq \mathfrak{B} \). Show that for each \( n < \omega \), there is a structure \( \mathfrak{C} \) such that

\[ \mathfrak{B} \prec_{n} \mathfrak{C} \quad \mathfrak{A} \prec \mathfrak{C} \]

if and only if \( \mathfrak{A} \prec_{n+1} \mathfrak{B} \).

3.8 Let \( T \) be a theory and let \( \phi, \psi \) be a pair of formulas. Show that the following are equivalent.

(i) We have

\[ T \vdash \phi \rightarrow \theta \quad T \vdash \theta \rightarrow \psi \]

for some \( \exists_1 \)-formula \( \theta \) with \( \partial \theta \subseteq \partial \phi \cup \partial \psi \).

(ii) We have

\[ \mathfrak{A} \models \phi(a) \implies \mathfrak{B} \models \psi(a) \]

for all models \( \mathfrak{A}, \mathfrak{B} \) of \( T \) with \( \mathfrak{A} \subseteq \mathfrak{B} \) and all points \( a \) of \( \mathfrak{A} \) matching the batch \( \partial \phi \cup \partial \psi \).
3.4 Directed families of structures

In this section we look at a method of combining many different structures into one structure. We then use this construction to obtain further results on the axiomatizability of theories.

3.20 DEFINITION. Let $\mathcal{A}$ be a non-empty family of structures (for a common language).

(a) The family is directed if for each $\mathcal{A}, \mathcal{B} \in \mathcal{A}$, there is some $\mathcal{C} \in \mathcal{A}$ with both $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{B} \subseteq \mathcal{C}$.

(b) The family is a chain if for each $\mathcal{A}, \mathcal{B} \in \mathcal{A}$, either $\mathcal{A} \subseteq \mathcal{B}$ or $\mathcal{B} \subseteq \mathcal{A}$.

(c) The family is a $\lambda$-chain (for some ordinal $\lambda$) if
\[ \mathcal{A} = \{ \mathcal{A}_i | i < \lambda \} \]
where $\mathcal{A}_i \subseteq \mathcal{A}_j$ for all $i \leq j < \lambda$.

(d) The family is an $\omega$-chain if
\[ \mathcal{A} = \{ \mathcal{A}_i | i < \omega \} \]
where $\mathcal{A}_i \subseteq \mathcal{A}_j$ for all $i \leq j < \omega$. ■

Clearly, every $\omega$-chain is an example of a $\lambda$-chain (for the case $\lambda = \omega$), and every $\lambda$-chain is a chain. A few moment's thought shows that every chain is directed.

Our first job is to show that for each directed family $\mathcal{A}$ we can construct a structure $\bigcup \mathcal{A}$ which extends each member of $\mathcal{A}$, and is as small as possible. Thus we produce a certain set $U$ and then furnish this to obtain a structure $U$ such that $\mathcal{A} \subseteq U$ for each $\mathcal{A} \in \mathcal{A}$.

3.21 CONSTRUCTION. Let $\mathcal{A}$ be a directed family of structures, as above. We produce $U = \bigcup \mathcal{A}$ as follows.

- Consider first the carriers of the members of $\mathcal{A}$. Let $U$ be the union of all these sets. Thus each $a \in U$ is also a member of at least one $\mathcal{A} \in \mathcal{A}$. It may be in at least two members, say $\mathcal{A}$ and $\mathcal{B}$, where neither of these extends the other. However, since $\mathcal{A}$ is directed, there is some $\mathcal{C} \in \mathcal{A}$ which extends both $\mathcal{A}$ and $\mathcal{B}$, and then $a$ is a member of $\mathcal{C}$. We will use this trick several times.

- Consider any constant $K$ of the language. Each $\mathcal{A} \in \mathcal{A}$ has a canonical interpretation $a = \mathcal{A}[K]$ of this constant, to give a member of $U$. How do these various members compare? Consider two members $\mathcal{A}, \mathcal{B}$ of $\mathcal{A}$ with interpretations $a = \mathcal{A}[K]$ and $b = \mathcal{B}[K]$ of the constant. We know there is some $\mathcal{C} \in \mathcal{A}$ which extends both $\mathcal{A}$ and $\mathcal{B}$, and then $a$ is a member of $\mathcal{C}$. We will use this trick several times.

Thus, working in $\mathcal{C}$ we have
\[ a = \mathcal{A}[K] = \mathcal{C}[K] = \mathcal{B}[K] = b \]
to show that $a, b$ are the same element of $U$. In other words, this constant $K$ determines a unique element of $U$. We let this be the interpretation of $K$. 
3. Basic methods

• Consider any \( n \)-placed relation symbol \( R \) if the language. We must decide whether or not \( [R]a_1 \cdots a_n \) is true for each \( a_1, \ldots, a_n \in U \).

Each \( a_i \) is an element of at least one member \( \mathfrak{A}_i \) of \( \mathcal{A} \). By repeated use of the directedness we see there is some \( \mathfrak{A} \in \mathcal{A} \) which extends each of these \( \mathfrak{A}_i \). Thus all of \( a_1, \ldots, a_n \) are members of some single \( \mathfrak{A} \in \mathcal{A} \). In that structure \( \mathfrak{A}[R]a_1 \cdots a_n \) is either true or false. However, as yet this truth value may depend on the holding structure \( \mathfrak{A} \) used.

Consider any \( \mathfrak{A}, \mathfrak{B} \in \mathcal{A} \), both of which contain \( a_1, \ldots, a_n \). There is some \( \mathfrak{C} \in \mathcal{A} \) which extends \( \mathfrak{A} \) and \( \mathfrak{B} \). Then

\[ \mathfrak{A}[R]a_1 \cdots a_n = \mathfrak{C}[R]a_1 \cdots a_n = \mathfrak{B}[R]a_1 \cdots a_n \]

to show that \( [R]a_1 \cdots a_n \) is independent of the choice of holding structure. This gives a single truth value for \( [R]a_1 \cdots a_n \).

• Consider any \( n \)-placed operation symbol \( O \) if the language. We must assign a value to \( [O]a_1 \cdots a_n \) for each \( a_1, \ldots, a_n \in U \). As in the previous case we can find a common holding structure \( \mathfrak{A} \in \mathcal{A} \) which contains each of \( a_1, \ldots, a_n \). This structure gives a value \( \mathfrak{A}[O]a_1 \cdots a_n \) in \( \mathfrak{A} \), but we need to check that this is independent of the choice of the holding structure \( \mathfrak{A} \). The argument for this is similar to the relation case.

This structure \( \mathfrak{U} = \bigcup \mathcal{A} \) is the union of the directed system \( \mathcal{A} \). Notice that we have \( \mathfrak{A} \subseteq \mathfrak{U} \) for each \( \mathfrak{A} \in \mathcal{A} \).

The construction \( \mathfrak{A} \rightarrow \bigcup \mathcal{A} \) has some preservation properties.

3.22 LEMMA. Let \( \mathcal{A} \) be a directed family of structures with union \( \mathfrak{U} \). Then

\[ \mathfrak{U} \models \text{Th}(\mathcal{A}) \cap \forall_2 \]

holds.

Proof. Consider any \( \forall_2 \)-sentence \( \sigma \) with \( \mathcal{A} \models \sigma \). We must show that \( \mathfrak{U} \models \sigma \).

We have

\[ \sigma = (\forall u)\phi(u) \]

for some \( \exists_1 \)-formula \( \phi(u) \) and batch \( u \) of variables. Consider any point \( a \) of \( \mathfrak{U} \) which matches \( u \). We must show that \( \mathfrak{U} \models \phi(a) \). Since \( \mathcal{A} \) is directed there is at least one \( \mathfrak{A} \in \mathcal{A} \) which contains all of \( a \). But \( \mathfrak{A} \models \sigma \) (by the choice of \( \sigma \)), and hence \( \mathfrak{A} \models \phi(a) \). Also, \( \mathfrak{A} \subseteq \mathfrak{U} \), and hence, since \( \phi \) is \( \exists_1 \), we have \( \mathfrak{U} \models \phi(a) \), as required.

Within a directed system \( \mathcal{A} \) there are many inclusions between its members. Some of these may have stronger preservation properties. In an extreme case each of these inclusions may be elementary.

3.23 LEMMA. Let \( \mathcal{A} \) be a directed family of structures with union \( \mathfrak{U} \). Suppose

\[ \mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec \mathfrak{B} \]

for all \( \mathfrak{A}, \mathfrak{B} \in \mathcal{A} \). Then

\[ \mathfrak{A} \prec \mathfrak{U} \]

for all \( \mathfrak{A} \in \mathcal{A} \).
3.4. Directed families of structures

Proof. We show the following by induction on \( n \).

[n] For each \( \forall_{2n} \)-formula \( \phi(u) \) and each \( \mathfrak{A} \in \mathcal{A} \) we have

\[
\mathfrak{A} \models \phi(a) \implies \mathfrak{U} \models \phi(a)
\]

for all points \( a \) of \( \mathfrak{A} \) (matching the batch of free variables \( u \) of \( \phi \)).

The base case, \( n = 0 \), is trivial.

For the induction step, \( n \mapsto n + 1 \), consider any \( \forall_{2n+2} \)-formula \( \phi(u) \). We know that \( \phi \) is

\[(\forall v)(\exists w)\psi(w, v, u)\]

for some \( \forall_{2n} \)-formula \( \phi(w, v, u) \). Consider any \( \mathfrak{A} \in \mathcal{A} \) and any point \( a \) of \( \mathfrak{A} \) for which

\[\mathfrak{A} \models \phi(a)\]

holds. We must show that

\[\mathfrak{U} \models \phi(a)\]

holds. To this end, consider any point \( b \) of \( \mathfrak{U} \) which matches \( v \), we must produce some point \( c \) of \( \mathfrak{U} \) which matches \( w \) for which

\[\mathfrak{U} \models \psi(c, b, a)\]

holds.

The point \( b \) comes from \( \mathfrak{B} \in \mathcal{A} \), and then there is some \( \mathfrak{C} \in \mathcal{A} \) which extends both \( \mathfrak{A} \) and \( \mathfrak{B} \), so that both \( a \) and \( b \) are points of \( \mathfrak{C} \). We have

\[\mathfrak{A} \models \phi(a) \quad \mathfrak{A} \prec \mathfrak{C}\]

so that \( \mathfrak{C} \models \phi(a) \), and hence \( \mathfrak{C} \models (\exists w)\psi(w, b, a) \) holds. This gives some point \( c \) of \( \mathfrak{C} \) with \( \mathfrak{C} \models \psi(c, b, a) \). The induction hypothesis \([n]\) now gives \( \mathfrak{U} \models \psi(c, b, a) \), as required. ■

With this result we can obtain a characterization of \( \forall_2 \)- axiomatizability that is rather different from the earlier characterization.

3.24 THEOREM. For each theory \( T \) (in some language \( L \)) the following are equivalent.

(i) \( T \) is \( \forall_2 \)-axiomatizable.

(ii) \( \mathcal{M} \mathcal{D}(T) \) is closed under unions of directed systems.

(iii) \( \mathcal{M} \mathcal{D}(T) \) is closed under unions of \( \omega \)-chains.

Proof. (i) \( \implies \) (ii). This follows by Lemma 3.22.

(ii) \( \implies \) (iii). This is trivial.

(iii) \( \implies \) (i). Assuming (iii), consider any model \( \mathfrak{A} \models T \cap \forall_2 \). It is sufficient to show that \( \mathfrak{A} \models T \). By Exercise 3.6 (which generalizes Lemma 3.16) we have \( \mathfrak{A} \prec_1 \mathfrak{B} \) for some \( \mathfrak{B} \models T \). Thus, by Lemma 3.19 we have

\[\mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{A}' \quad \mathfrak{B} \models T \quad \mathfrak{A} \prec \mathfrak{A}'\]
for some structures $\mathfrak{B}, \mathfrak{A}$. Notice that $\mathfrak{A} \models T \cap \forall_2$.

By iterating this construction we obtain two $\omega$-chains

$$A = \{A_i \mid i < \omega\} \quad B = \{B_i \mid i < \omega\}$$

where $A_0 = A$ and

$$A_i \subseteq B_i \subseteq A_{i+1} \quad B_i \models T \quad A_i \prec A_{i+1}$$

for each $i < \omega$.

Each of these chains has a union. But the two chains interlace, so there is a single structure $\mathfrak{U}$ which is the union of both $A$ and $B$. By construction, $B$ is a chain of models of $T$, and hence (iii) gives $\mathfrak{U} \models T$. Again by construction, $A$ is a chain of elementary embedding, and hence Lemma 3.23 gives

$$\mathfrak{A} \prec \mathfrak{U} \models T$$

so that $\mathfrak{A} \models T$, as required. ■

To conclude this section we look at a situation which is slightly unusual in model theory. We look at intersections of structures.

Given any family $\mathcal{A}$ is structures (for some language) we can make sense of the intersection $\bigcap \mathcal{A}$. This is either empty, or a substructure of each of the members of $\mathcal{A}$. Note also, that if the underlying language contains at least one constant symbol, then $\bigcap \mathcal{A}$ can not be empty.

We show how to exhibit the union of a directed family of models as a binary intersection. This uses a different kind of enrichment of the underlying language.

3.25 LEMMA. Let $T$ be a theory (in some language $L$). Let $\mathcal{A}$ be a directed family of models of $T$, and let $\mathfrak{U} = \bigcup \mathcal{A}$. Then there are models $\mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ of $T$ such that

$$\mathfrak{U} \subseteq \mathfrak{B} \prec \mathfrak{D} \quad \mathfrak{U} \subseteq \mathfrak{C} \subseteq \mathfrak{D} \quad \mathfrak{U} = \mathfrak{B} \cap \mathfrak{C}$$

hold.

Proof. By Lemma 3.22 we have $\mathfrak{U} \models T \cap \forall_2$, and then Exercise 3.6 gives some model $\mathfrak{B}$ with $\mathfrak{U} \prec_1 \mathfrak{B} \models T$. Our problem is to construct $\mathfrak{C}$ and $\mathfrak{D}$.

Let $b$ be an enumeration of $\mathfrak{B}$, and consider $Th(\mathfrak{B}, b)$ in the enriched language. We now form a further enrichment by adding a new 1-placed relation symbol $R$. Within the language $L + \{R\}$ let $\Gamma$ be the set of sentences

$$R \text{ is the carrier of a substructure (of the parent structure),}$$

and this substructure is a model of $T$

(which isn’t too hard to formalize). Using this consider the set

$$Th(\mathfrak{B}, b) \cup \Gamma \cup \{Ra \mid a \text{ from } \mathfrak{U}\} \cup \{\neg Rb \mid b \text{ from } \mathfrak{B} \text{ but not from } \mathfrak{U}\}$$

of sentences of the largest language. Any finite subset of this is a subset of

$$Th(\mathfrak{B}, b) \cup \Gamma \cup \{Ra_1, \ldots, Ra_m\} \cup \{\neg Rb \mid b \text{ from } \mathfrak{B} \text{ but not from } \mathfrak{U}\}$$
for some elements \(a_1, \ldots, a_m\) of \(U\). These elements all belong to some \(A \in A\). Setting \(R = A\) (the carrier of \(A\)) produces a model of the smaller set of sentences. Thus the larger set is consistent.

Let \((\mathcal{D}, C, b)\) be any model of this larger set. From \(Th(\mathcal{B}, b)\) we have \(\mathcal{B} \prec \mathcal{D}\). From \(\Gamma\) the subset \(C\) is the carrier of a substructure \(\mathcal{C} \subseteq \mathcal{D}\) which is a model of \(T\). From the third component we have \(U \subseteq \mathcal{C}\), and the fourth component gives \(\mathcal{B} \cap C \subseteq U\), which is what we want. ■

Occasionally we meet a theory of the following kind.

3.26 DEFINITION. A theory \(T\) is convex if

\[ A \cap B \models T \]

for all models \(A, B, C\) of \(T\) with \(A \subseteq C\), and \(B \subseteq C\), and \(A \cap B \neq \emptyset\). ■

The final result of this section is rather surprising.

3.27 THEOREM. Each convex theory \(T\) is \(\forall_2\)-axiomatizable.

Proof. By Lemma 3.25 the class \(Md(T)\) is closed under unions of directed families. ■

We will use directed families of structures many times.

Exercises

3.9 Let \(A\) be a directed family of structures with union \(U\). Show that if, for some \(n < \omega\),

\[ A \subseteq B \implies A \prec_n B \]

for all \(A, B \in A\), then

\[ A \prec_n U \]

for all \(A \in A\).

3.10 Let \(A\) be a directed family of structures with union \(U\). Suppose that for each pair \(A \subseteq B\) of members of \(A\), there is some \(B \subseteq C \in A\) with \(A \prec C\). Show that, under these circumstances, \(A \prec U\) for each \(A \in A\).

3.11 Two directed families \(A\) and \(B\) of structures interlace if both

- For each \(A \in A\), there is some \(B \in B\) with \(A \subseteq B\)
- For each \(B \in B\), there is some \(A \in A\) with \(B \subseteq A\)

hold. Show that

\[ \bigcup A = \bigcup B \]

for such families.

3.12 Let \(A\) be a directed family of models of some theory \(T\), and let \(U\) be \(\bigcup A\). Show there are models \(\mathcal{B}, \mathcal{C}, \mathcal{E}\) of \(T\) with

\[ U \subseteq \mathcal{B} \subseteq \mathcal{E} \quad \mathcal{U} \subseteq \mathcal{C} \subseteq \mathcal{E} \quad \mathcal{U} = \mathcal{B} \cap \mathcal{C} \]

and with \(\mathcal{U} \prec \mathcal{E}\).
3.5 The up and down techniques

In this section we obtain two consequences of refined compactness, Theorem 1.27, which are used all the time in model theory, often without even mentioning it. The results allow us to move between cardinalities, provided we stay above the cardinality of the underlying language.

3.28 THEOREM. Let \( L \) be a language, let \( \mathfrak{A} \) be a \( L \)-structure, and let \( \kappa_d, \kappa_u \) be cardinals such that
\[
|L| \leq \kappa_d \leq |\mathfrak{A}| \leq \kappa_u
\]
hold. Then there are structures \( \mathfrak{A}_d, \mathfrak{A}_u \) (for \( L \)) such that
\[
|\mathfrak{A}_d| = \kappa_d \quad \mathfrak{A}_d \prec \mathfrak{A} \prec \mathfrak{A}_u \quad |\mathfrak{A}_u| = \kappa_u
\]
hold.

These are often referred to as the
downward Löwenheim-Skolem \quad upward Löwenheim-Skolem
results, respectively, but that can be a bit of a mouthful.

How can we produce these two structures \( \mathfrak{A}_d \) and \( \mathfrak{A}_u \)? Surprisingly, the larger one is easier to construct.

The up technique

Given a structure \( \mathfrak{A} \) for a language \( L \) and a cardinal \( \kappa \), how can we produce a structure \( \mathfrak{A} \prec \mathfrak{B} \) with \( |\mathfrak{B}| = \kappa \)? Clearly, to do this we need \( |\mathfrak{A}| \leq |\mathfrak{B}| = \kappa \), but there are other restrictions as well.

We use the elementary (full) version of the diagram technique.

Let \( \mathfrak{a} \) be an enumeration of the whole of \( \mathfrak{A} \). Let \( \mathfrak{b} \) be an enumeration of a further \( \kappa \) new parameters. Unlike the members of \( \mathfrak{a} \), these further parameters are not associated with any elements of \( \mathfrak{A} \). We work in the enrichment \( L(\mathfrak{a}, \mathfrak{b}) \) of \( L \). Note that
\[
|L(\mathfrak{a}, \mathfrak{b})| = |L| + |\mathfrak{a}| + |\mathfrak{b}| = |L| + \kappa
\]
since \( |\mathfrak{a}| = |\mathfrak{A}| \leq \kappa = |\mathfrak{b}| \). We assume \( |L| \leq \kappa \) to get \( |L(\mathfrak{a}, \mathfrak{b})| = \kappa \).

Look at the set of sentences
\[
Th(\mathfrak{A}, \mathfrak{a}) \cup \text{‘the members of } \mathfrak{b} \text{ are pairwise distinct’}
\]
of \( L(\mathfrak{a}, \mathfrak{b}) \). Observe that this set is finitely satisfiable in \( \mathfrak{A} \). You should check the details of this and note that only
\[
\aleph_0 \leq |\mathfrak{A}|
\]
is needed.

Refined compactness gives a model \( (\mathfrak{B}, \mathfrak{a}, \mathfrak{b}) \) of this set where \( |\mathfrak{B}| = |L(\mathfrak{a}, \mathfrak{b})| = \kappa \). The first component ensures that \( \mathfrak{A} \prec \mathfrak{B} \), which is what we want.

You should observe the potential of this technique. In fact, we have already proved more than was asked for.
3.29 **THEOREM.** Let $L$ be a language, let $A$ be in infinite $L$-structure, and let $\kappa$ be a cardinal with $|L| \leq \kappa$, $|A| \leq \kappa$. Then there is a structure $\mathcal{B}$ (for $L$) such that $A \prec \mathcal{B}$ and $|\mathcal{B}| = \kappa$.

As a further refinement we might want to arrange that the structure $\mathcal{B}$ has several other properties. Maybe this can be done by controlling these properties by sentences of the enriched language.

**The down technique**

Given a structure $A$ for a language $L$ and a cardinal $\kappa$, how can we produce a structure $\mathcal{B} \prec A$ with $|\mathcal{B}| = \kappa$? Clearly, to do this we need $\kappa = |\mathcal{B}| \leq |A|$, but there are other restrictions as well.

To produce $\mathcal{B}$ we work in an enrichment of $L$, but this time we adjoin certain new operation symbols.

Recall that a subset of $B$ of the carrier $A$ of $A$ is the carrier of a substructure $\mathcal{B} \subseteq A$ if and only if $B$ is closed under the distinguished attributes of $A$. In general, this closure does not ensure $\mathcal{B} \prec A$. However it does if $Th(A)$ has $EQ$, so we attempt to arrange this.

We ‘skolemize’ the language by adjoining certain skolem operation symbols (which are then interpreted as skolem functions).

Consider an arbitrary formula $\phi(v,u)$ of the language $L$, where $u$ is a batch of $n$ variables and $v$ is a single nominated target variable. For each such formula and nomination we select a new $n$-placed operation symbol $f = f_{\phi}$, and look at the sentence

$$Skol(\phi)(\forall u)[(\exists v)\phi(v,u) \rightarrow \phi(fu,u)]$$

in the enriched language. We do this for each formula $\phi$ and nominated target variable $v$, using a different skolem operation symbol $f$ for each such pair. Let $L'$ be the enrichment of $L$ obtained by adjoining each of the skolem operation symbols. Notice that $|L'| = |L|$, so that although $L'$ has many more formulas, its cardinality is the same as $L$.

Let $S'$ be the theory in $L'$ axiomatized by the set of all $Skol(\phi)$.

Consider any $L$-structure $\mathcal{A}$. We show how to enrich this to a $L'$-structure by adjoining an interpretation $\mathcal{A}'[f]$ for each skolem operation symbol $f$. This, of course, will be a skolem function.

Let $\phi(v,u)$ be the formula associated with $f$. Consider any point $a$ of $\mathcal{A}$ which matches $u$. We require a value $\mathcal{A}'[f]a$. There are two cases to consider.

Suppose $\mathcal{A} \models (\exists v)\phi(v,a)$. Then $\mathcal{A} \models \phi(b,a)$ for some element $b$ of $\mathcal{A}$, and we let $\mathcal{A}'[f]a$ be any such element.

Suppose $\mathcal{A} \models \neg(\exists v)\phi(v,a)$. Then we let $\mathcal{A}'[f]a$ be any element whatsoever.

This ensures that $\mathcal{A}' \models Skol(\phi)$. In fact

$$\mathcal{A}' \models (\forall u)[(\exists v)\phi(v,u) \leftrightarrow \phi(fu,u)]$$

holds.

At first sight it seems that $Th(\mathcal{A}')$ has $EQ$ (for, as we have just seen, the operation symbol $f$ can be used to eliminate certain quantifications $(\exists v)$). Another look reveals
an error. In Th(\(\mathfrak{A}'\)) we can eliminated certain quantifiers, but only from formulas of the original language \(L\). As yet, we can not eliminate quantifiers from all formulas of \(L'\), for such a formula may not have an associated skolem operation symbol. This is easy to correct.

We iterate the process

\[ L \subseteq L' \subseteq L' \subseteq \cdots \subseteq L^{(r)} \subseteq \cdots \]

to produce an ascending \(\omega\)-chain of languages each of cardinality \(|L|\). Let \(L^*\) be the union of these languages. Thus \(|L^*| = |L|\).

Each \(L\)-structure \(\mathfrak{A}\) can be enriched, step by step, to a \(L^*\)-structure \(\mathfrak{A}^*\). Furthermore, by accumulation, we see that the following holds.

For each formula \(\phi(v,u)\) of \(L^*\) where \(v\) is a single variable, there is an operation symbol \(f\) of \(L^*\) such that

\[ \mathfrak{A}^* \models (\forall u)[(\exists v)\phi(v,u) \leftrightarrow \phi(fu,u)] \]

holds.

(The formula \(\phi\) lives in some \(L^{(r)}\), and then \(f\) is in \(L^{(r+1)}\).) This ensures that \(Th(\mathfrak{A}^*)\) has \(EQ\).

With this we can produce the required substructure of \(\mathfrak{A}\).

Suppose \(|L| \leq \kappa \leq |\mathfrak{A}|\). Let \(\mathfrak{A}^*\) be a skolem enrichment of \(\mathfrak{A}\), as described above. Take any subset \(X \subseteq A\) (the carrier of \(\mathfrak{A}\)) with \(|X| \leq \kappa\). Let \(B\) be the closure of \(X\) in \(A\) under all the distinguished attributes of \(\mathfrak{A}^*\). Note that \(|B| = \kappa\). Furthermore, \(B\) is the carrier of a substructure \(\mathfrak{B}^* \subseteq \mathfrak{A}^*\). Even more, since \(Th(\mathfrak{A}^*)\) has \(EQ\), we have \(\mathfrak{B}^* \prec \mathfrak{A}^*\). Let \(\mathfrak{B}\) be the \(L\)-structure obtained from \(\mathfrak{B}^*\) by removing the extra skolem operations. We have \(\mathfrak{B} \prec \mathfrak{A}\) and \(|\mathfrak{B}| = \kappa\), as required.

Again this technique has quite a bit of potential. We have already proved the following.

3.30 THEOREM. Let \(L\) be a language, let \(\mathfrak{A}\) be a \(L\)-structure, and let \(\kappa\) be a cardinal with \(|L| \leq \kappa \leq |\mathfrak{A}|\). for each subset \(X\) of \(\mathfrak{A}\) with \(|X| \leq \kappa\), there is a structure \(\mathfrak{B} \prec \mathfrak{A}\) with \(|\mathfrak{B}| \leq \kappa\) such that \(X\) is a subset of \(\mathfrak{B}\).

Perhaps, with some effort, we can make these skolem functions do other things.

Exercises

3.13 Let \(\mathfrak{A} \subseteq \mathfrak{B}\) be a pair of structure and suppose the following holds. For each formula \(\phi(v,u)\) with a nominated variable \(v\), if \(\mathfrak{B} \models (\exists v)\phi(v,a)\) for some point \(a\) from \(\mathfrak{A}\), then \(\mathfrak{B} \models \phi(b,a)\) for some element \(b\) from \(\mathfrak{A}\) (not just \(\mathfrak{B}\)). Show that \(\mathfrak{A} \prec \mathfrak{B}\).
4

Model complete and submodel complete theories

In Definition 2.1 we introduced the notion of a theory $T$ having $EQ$ (elimination of quantifiers). This looks like a rather syntactic idea. In Sections 2.3 and 2.2 we gave two examples of such theories. The methods of proof used there look rather intricate, and special to those particular theories. In this chapter we will show there is a rather general condition which ensures that a theory has $EQ$. Later, in Chapter 6, we will show how these conditions can be verified by an analysis of the spectrum of models of a theory.

4.1 Model complete theories

Recall that by Definition 3.10 and Lemma 3.12, for a given a theory $T$ and a submodel $\mathfrak{A} \in S(T)$ we can form a consistent set

$$T[\mathfrak{A}] = T \cup \text{Diag}(\mathfrak{A}, a)$$

of sentences in the $\mathfrak{A}$-enriched language $L(\mathfrak{A})$ of the parent language $L$. This set $T[\mathfrak{A}]$ generates a (consistent) $L(\mathfrak{A})$-theory, and we can ask that this enrichment of $T$ has a certain property. We can ask for such a property for a whole family of submodels.

4.1 DEFINITION. A theory $T$ is model complete if the enriched theory (axiomatized by) $T[\mathfrak{A}]$ is complete for each $\mathfrak{A} \in \mathcal{Md}(T)$. ■

This is the original definition of model completeness but, as the next result shows, the notion can be characterized in a rather more amenable fashion.

4.2 LEMMA. A theory $T$ is model complete precisely when

$$\mathfrak{A} \subseteq \mathfrak{B} \Rightarrow \mathfrak{A} \prec \mathfrak{B}$$

for each pair $\mathfrak{A}, \mathfrak{B}$ of models of $T$.

Almost invariably it is this characterization that we use. Only occasionally do we return to the official Definition 4.1.

(You might want to think about what happens if we add to $T$ the diagram of a submodel of $T$, or a special kind of submodel. We will return to this idea several times.)

Model completeness is one of the most important notions of model theory. Here is why.

4.3 THEOREM. For each theory $T$ the following are equivalent.

(i) Each formula is $T$-equivalent to an $\forall_1$-formula.
(ii) Each formula is $T$-equivalent to an $\exists_1$-formula.

(iii) Each $\forall_1$-formula is $T$-equivalent to an $\exists_1$-formula.

(iv) We have

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec_1 \mathfrak{B}$$

for each pair $\mathfrak{A}, \mathfrak{B}$ of models of $T$.

(v) $T$ is model complete.

**Proof.** The implication $(i) \implies (ii)$ holds by a simple use of negated formulas.

$(ii) \implies (iii)$. This is trivial.

$(iii) \implies (iv)$. This is immediate.

$(iv) \implies (v)$. Suppose $(iv)$ holds. For each models $\mathfrak{A} \subseteq \mathfrak{B}$ we have $\mathfrak{A} \prec_1 \mathfrak{B}$ and hence, by Lemma 3.19, we have

$$\mathfrak{A} \prec \mathfrak{A}' \quad \mathfrak{B} \subseteq \mathfrak{A}'$$

for some model $\mathfrak{A}'$ of $T$. But now we have $\mathfrak{B} \prec_1 \mathfrak{A}'$ so that a repeat of the argument gives

$$\mathfrak{B} \prec \mathfrak{B}' \quad \mathfrak{A}' \subseteq \mathfrak{B}'$$

for some model $\mathfrak{B}'$ of $T$. By iterating this construction we produce two interlacing elementary chains

$$\mathcal{A} = \{ \mathfrak{A}_i \mid i < \omega \} \quad \mathcal{B} = \{ \mathfrak{B}_i \mid i < \omega \}$$

with

$$\mathfrak{A}_i \prec \mathfrak{A}_{i+1} \quad \mathfrak{A}_i \subseteq \mathfrak{B}_i \subseteq \mathfrak{A}_{i+1} \subseteq \mathfrak{B}_{i+1} \quad \mathfrak{B}_i \prec \mathfrak{B}_{i+1}$$

for each $i < \omega$.

Since the chains interlace, they have a common union $\mathcal{U}$.

Since the chains are elementary we have

$$\mathfrak{A} = \mathfrak{A}_0 \prec \mathcal{U} \quad \mathfrak{B} = \mathfrak{B}_0 \prec \mathcal{U}$$

and hence $\mathfrak{A} \prec \mathfrak{B}$, as required.

$(v) \implies (i)$. Consider any formula $\phi(v)$ where $v$ is the batch of free variables of $\phi$. We enrich the underlying language with parameters $a$ matching the list $v$. Within this enriched language let $\Psi(a)$ be the set of $\forall_1$-sentences $\psi(a)$ such that

$$T \vdash \phi(a) \implies \psi(a)$$

holds. If $T(a)$ is the theory in the enriched language axiomatized by $T \cup \{ \phi(a) \}$, then $\Psi(a) = T(a) \cap \forall_1$. It suffices to show that

$$T \cup \Psi(a) \vdash \phi(a)$$

holds.

Consider any model $\langle \mathfrak{A}, a \rangle$ of $T \cup \Psi(a)$. By Lemma 3.16, there is some model $\langle \mathfrak{B}, a \rangle$ of $T(a)$ with $\langle \mathfrak{A}, a \rangle \subseteq \langle \mathfrak{B}, a \rangle$. But now

$$\mathfrak{A} \models T \quad \mathfrak{A} \subseteq \mathfrak{B} \quad \mathfrak{B} \models T \quad \mathfrak{B} \models \phi(a)$$

so that, by the assumption $(iv)$, we have $\mathfrak{A} \prec \mathfrak{B}$ and hence $\mathfrak{A} \models \phi(a)$, which leads to the required result. $\blacksquare$

This result indicates that model completeness is a weakened form of quantifiers elimination. Thus the property $EQ$ must be model completeness together with some other property. What can this extra property be?
Exercises

4.1 Prove Lemma 4.2.

4.2 Consider the proof of the implication \((iv) \Rightarrow (v)\) of Theorem 4.3. That uses the interlacing chains argument. Here is another way of obtaining the implication.

Show that if a theory \(T\) satisfies

\[
[0] \ A \subseteq B \implies A \prec_1 B
\]

for all models \(A, B\), then for each \(n \in \mathbb{N}\) it also satisfies

\[
[n] \ A \prec_n B \implies A \prec_{n+1} B
\]

for all models \(A, B\), and hence is model complete.

4.3 (a) Show that if a theory has \(E_Q\) then it is model complete.

(b) Show that a model complete theory is \(\forall_2\)-axiomatizable.

4.2 Two structural properties

In this section we look at two structural properties each of which the class of models of a theory may or may not have. At first sight you may wonder why these two properties are being analysed here. We will see why shortly.

4.4 Definition. (jep) A theory \(T\) has JEP (the joint embedding property) if for each pair \(A, B\) of models of \(T\), there is a wedge of embeddings

\[
\begin{array}{ccc}
& C & \\
A & \nearrow & B
\end{array}
\]

to some model \(C\) of \(T\).

(ap) A theory \(T\) has AP (the amalgamation property) if for each wedge of embedding between models of \(T\), as to the left

\[
\begin{array}{ccc}
B & C & D \\
A & \nearrow & A
\end{array}
\]

there is a model \(D\) of \(T\) and a commuting square of embeddings, as to the right.

These two properties are independent. A theory may have one without the other. Of course, a theory may have both, or it may have neither. If you know a bit of category theory (and you should) you may misread AP. All that is required is that the square of embeddings does commute. It need not have any universal properties, so any such square is as good as any other.

Why are we looking at these two properties here? We have to look at them sometime, and because of their similarities it makes sense to look at them together. Also JEP is
4. Model complete and submodel complete theories

slightly easier to analyse then $AP$, so it makes sense to look at $JEP$ just before $AP$. But why do that here? As we will see in the next section the structural property $AP$ is involved with both model completeness and elimination of quantifiers.

In this section we produce syntactic characterizations of both $JEP$ and $AP$. To do that we use the method of diagrams set up in Chapter 3. However, the proofs are slightly more involved.

The structural property $JEP$ of the class of models of a theory corresponds to a certain primeness property of the corresponding consequence relation.

4.5 THEOREM. A (consistent) theory $T$ has $JEP$ precisely when

$$T \vdash \alpha \lor \beta \implies T \vdash \alpha \text{ or } T \vdash \beta$$

for each pair $\alpha, \beta$ of $\forall_1$-sentences (of the underlying language).

Proof. Suppose that $T$ has the primeness property, consider any pair $\mathfrak{A}, \mathfrak{B}$ of models of $T$. Let $a$ be a enumeration of the whole of $\mathfrak{A}$, and let $b$ be a enumeration of the whole $\mathfrak{B}$. Add these enumeration to the underlying language $L$ to form a doubly enriched language $L(a, b)$. Thus $L(a, b)$ is the amalgam of the two languages $L(\mathfrak{A})$ and $L(\mathfrak{B})$. We may assume that the two sets of new parameters are disjoint. To produce the joint embedding it suffices to show that

$$T \cup \text{Diag}(\mathfrak{A}, a) \cup \text{Diag}(\mathfrak{B}, b)$$

is consistent.

By way of contradiction, suppose this set is not consistent. Then, by compactness, it has a finite subset which is inconsistent. This gives a sentence $\gamma(a)$ from $\text{Diag}(\mathfrak{A}, a)$ and a sentence $\delta(b)$ from $\text{Diag}(\mathfrak{B}, b)$ such that

$$T \cup \{\gamma(a), \delta(b)\}$$

is inconsistent. Thus we have

$$\mathfrak{A} \models \gamma(a) \quad \mathfrak{B} \models \delta(b) \quad T \vdash \neg \gamma(a) \lor \neg \delta(b)$$

for some quantifier-free formulas $\gamma(v), \delta(w)$ of the underlying language, and points $a$ from $\mathfrak{A}$ and $b$ from $\mathfrak{B}$ matching the batches of free variables $v$ and $w$, respectively.

Consider the third condition

$$T \vdash \neg \gamma(a) \lor \neg \delta(b)$$

of the triple. Since the constants $a$ and $b$ do not occur in $T$ we can treat them as variable. Thus we have

$$T \vdash \neg \gamma(v) \lor \neg \delta(w)$$

for disjoint batches $v, w$ of actual variables of the underlying language. Since $T$ is a set of sentences we see that

$$T \vdash (\forall v)\neg \gamma(v) \lor (\forall w)\neg \delta(w)$$

holds.

Both $\gamma$ and $\delta$ are quantifier-free formulas, so the assumed primeness property of $T$ ensures that one of

$$T \vdash (\forall v)\neg \gamma(v) \quad T \vdash (\forall w)\neg \delta(w)$$

holds.
holds. But $\mathfrak{A}, \mathfrak{B} \models T$ and both

$$\mathfrak{A} \models \gamma(a) \quad \mathfrak{B} \models \delta(b)$$

which is the required contradiction.

The proof of the converse implication is an easy exercise. □

We now do a similar job on $AP$. The characterization is again a primeness property, but this time it involves formulas not just sentences.

Remember that a set $\Gamma$ of formulas is a type if

$$\partial \Gamma = \bigcup \{ \partial \phi \mid \phi \in \Gamma \}$$

is finite. Given such a type we set

$$\neg \Gamma = \{ \neg \phi \mid \phi \in \Gamma \}$$

which, of course, is also a type with $\partial (\neg \Gamma) = \partial \Gamma$.

4.6 THEOREM. For each theory $T$ the following are equivalent.

(i) $T$ has $AP$.

(ii) For each pair $\psi, \phi$ of $\forall_1$-formulas with $T \vdash \psi \lor \phi$ we have $T \vdash \lambda \lor \rho$ $T \vdash \lambda \rightarrow \psi$ $T \vdash \rho \rightarrow \phi$ for some pair $\lambda, \rho$ of $\exists_1$-formulas.

Proof. (i)$\Rightarrow$(ii). Consider any $\forall_1$-formulas $\psi, \phi$ such that $T \vdash \psi \lor \phi$ holds. Let $v$ be the batch of free variables occurring in these formulas. Let $\Sigma(T, \psi)$ $\Sigma(T, \phi)$ be the respective sets of $\exists_1$-formulas

$$\lambda \quad \rho$$

such that $T \vdash \lambda \rightarrow \psi$ $T \vdash \rho \rightarrow \phi$ and $\partial \lambda \cup \partial \rho \subseteq v$.

Each of these sets is an $\exists_1$-type, and so each of the two sets of negations $\neg \Sigma(T, \psi)$ $\neg \Sigma(T, \phi)$
is an \( \forall_1 \)-type. We will show that

\[ T \cup \neg \Sigma(T, \psi) \cup \neg \Sigma(T, \phi) \]

is inconsistent.

How does this help? Observe that both \( \Sigma(T, \psi) \) and \( \Sigma(T, \phi) \) are closed under disjunctions. Thus, assuming the set above is inconsistent, we obtain \( \lambda \in \Sigma(T, \psi) \) and \( \rho \in \Sigma(T, \phi) \) such that \( T \cup \{ \neg \lambda, \neg \rho \} \) is inconsistent. In other words, \( T \vdash \lambda \lor \rho \), and we are done.

It remains to show that the displayed set is inconsistent. This is where the assumption (i) is needed.

We make a preliminary observation.

Suppose we have a model \( \mathfrak{A} \models T \) and a point \( a \) of \( \mathfrak{A} \) which realizes \( \neg \Sigma(T, \psi) \). A simple argument show that

\[ T \cup \text{Diag}(\mathfrak{A}, a) \cup \{ \neg \psi(a) \} \]

is consistent (where \( a \) is an enumeration of \( \mathfrak{A} \)). You should make sure you can produce this argument. A use of compactness gives us an embedding of \( \mathfrak{A} \) into some model \( \mathfrak{B} \models T \) where \( \mathfrak{B} \models \neg \psi(a) \).

With this we can complete the whole argument.

By way of contradiction, suppose the displayed set is consistent. Thus it is realized by some point \( a \) of some model \( \mathfrak{A} \) of \( T \). Using the two components we obtain a wedge of embeddings between models of \( T \), as in Definition 4.4(ap), where

\[ \mathfrak{B} \models \neg \psi(a) \quad \mathfrak{C} \models \neg \phi(a) \]

holds. Since \( T \) has \( AP \) (by (i)) this gives us a commuting square of embeddings between models of \( T \) with some model \( \mathfrak{D} \) as the apex. Both \( \neg \psi \) and \( \neg \phi \) are \( \exists_1 \)-formulas, and hence \( \mathfrak{D} \models \neg (\psi \lor \phi)(a) \). This is the contradiction (since \( T \vdash \psi \lor \phi \)).

(ii)\( \Rightarrow \) (i). Suppose (ii) holds and consider any wedge of embeddings between models of \( T \), as in Definition 4.4(ap). By replacing \( \mathfrak{B} \) and \( \mathfrak{C} \) by suitably isomorphic copies, we may suppose

\[ \mathfrak{A} \subseteq \mathfrak{B} \quad \mathfrak{A} \subseteq \mathfrak{C} \quad \mathfrak{B} \cap \mathfrak{C} = \mathfrak{A} \]

hold. Let \( a, b, c \) be enumerations of \( \mathfrak{A}, \mathfrak{B}, \mathfrak{C} \) respectively. It suffices to show that

\[ T \cup \text{Diag}(\mathfrak{B}, b, a) \cup \text{Diag}(\mathfrak{C}, c, a) \]

is consistent.

By way of contradiction, suppose this set is not consistent. Then we have quantifier-free formulas \( \beta(v, u), \gamma(w, u) \) (of the underlying language) and points \( a, b, c \) such that

\[ T \vdash \neg \beta(v, u) \lor \neg \gamma(w, u) \quad \mathfrak{B} \models \beta(b, a) \quad \mathfrak{C} \models \gamma(c, a) \]

hold. The lists of variables \( u, v, w \) are disjoint so we have

\[ T \vdash \psi \lor \phi \]

where

\[ \psi(u) = (\forall v) \neg \beta(v, u) \quad \phi(u) = (\forall w) \neg \gamma(w, u) \]
are the two components. Both of these are \( \forall_1 \)-formulas, so that (ii) gives
\[
T \vdash \lambda \lor \rho \quad T \vdash \lambda \rightarrow \psi \quad T \vdash \rho \rightarrow \phi
\]
for some \( \exists_1 \)-formulas \( \lambda, \rho \). The first of these gives \( \mathfrak{A} \models (\lambda \lor \rho)(a) \), and hence \( \mathfrak{A} \models \lambda(a) \), say. Since \( \mathfrak{A} \subseteq \mathfrak{B} \), this gives \( \mathfrak{B} \models \lambda(a) \), and hence \( \mathfrak{B} \models \psi(a) \) follows from the second. But now \( \mathfrak{B} \models \neg \beta(b, a) \), which is the required contradiction. \( \blacksquare \)

You will need time to digest this proof.

**Exercises**

4.4 Complete the proof of Theorem 4.5.

4.5 Show that a theory \( T \) has \( AP \) if and only if \( T \cap \forall_2 \) has \( AP \).

4.6 Show that each model complete theory has \( AP \).

4.7 Let \( T \) be a theory such that \( T \cap \forall_1 \) has \( AP \). Show that for each \( \exists_1 \)-formula \( \theta \) and \( \forall_1 \)-formula \( \phi \) with \( T \vdash \theta \rightarrow \phi \), there is a quantifier-free formula \( \delta \) (with the same free variables) such that both \( T \vdash \theta \rightarrow \delta \) and \( T \vdash \delta \rightarrow \phi \) hold.

4.8 The proof of Theorem 4.6 can be modified to produce a characterization of certain kinds of submodels of a theory.

A submodel \( \mathfrak{A} \) of a theory \( T \) is an **amalgamation base for \( T \)** if for each wedge of embedding from \( \mathfrak{A} \) to models of \( T \), as to the left of the table in Definition 4.4(ap), there is a model \( \mathfrak{D} \) of \( T \) and a commuting square of embeddings, as to the right of that table, closing the given embeddings. Thus \( T \) has \( AP \) exactly when each model is an amalgamation base for \( T \).

Show that a submodel \( \mathfrak{A} \) of \( T \) is an amalgamation base for \( T \) if and only if for each pair \( \psi, \phi \) of \( \forall_1 \)-formulas with
\[
T \vdash \psi \lor \phi
\]
the structure \( \mathfrak{A} \) omits the \( \forall_1 \)-type
\[
\neg \Sigma(T, \psi) \cup \neg \Sigma(T, \phi)
\]
(where these are as used in the proof of Theorem 4.6).

4.9 For each theory \( T \) let \( \mathcal{B}(T) \) be the class of amalgamation bases for \( T \). Show the following.

(a) \( \mathcal{B}(T) = \mathcal{A}(T \cap \forall_1) \).
(b) If \( \mathfrak{A} \prec_1 \mathfrak{A}' \) for some \( \mathfrak{A}' \in \mathcal{B}(T) \), then \( \mathfrak{A} \in \mathcal{A}(T) \).
(c) \( \mathcal{B}(T) \) is closed under unions of directed families.
4.3 Submodel complete theories

In Section 4.2 we claimed that AP had some connection with the model theoretic properties of a theory. Let’s begin this section with an illustration of that.

In the next result we show how, under appropriate conditions, a wedge, as on the left

\[ \begin{array}{c}
\mathcal{B} \\
\downarrow f \\
\mathcal{A} \\
\downarrow g \\
\mathcal{C}
\end{array} \]

can be completed to a square, as on the right.

4.7 LEMMA. Given a wedge, as to the left above, where \( f \) is a \( \prec_1 \)-embedding and \( g \) is an embedding, there is a commuting square, as to the right above where \( h \) is an embedding and \( k \) is an elementary embedding.

Proof. By replacing \( \mathcal{B} \) and \( \mathcal{C} \) by suitable isomorphic copies, we can suppose that

\[ \begin{array}{c}
\mathfrak{A} \prec_1 \mathcal{B} \quad \mathfrak{A} \subseteq \mathcal{C} \quad \mathcal{B} \cap \mathcal{C} = \mathfrak{A}
\end{array} \]

hold. Let \( a, b, c \) be enumerations of \( \mathfrak{A}, \mathcal{B}, \mathcal{C} \), respectively. It suffices to show that

\[ Th(\mathcal{C}, c, a) \cup Diag(\mathcal{B}, b, a) \]

is consistent.

If this is not consistent then there is some quantifier-free formula \( \delta(v, u) \) (of the underlying language) and appropriate points \( a, b \) such that

\[ Th(\mathcal{C}, c, a) \vdash \neg \delta(b, a) \quad \mathcal{B} \models \delta(b, a) \]

hold. The first of these gives

\[ Th(\mathcal{C}, c, a) \vdash (\forall v) \neg \delta(v, a) \]

and hence \( \mathcal{C} \models (\forall v) \neg \delta(v, a) \). But then \( \mathfrak{A} \models (\forall v) \neg \delta(v, a) \), so that \( \mathcal{B} \models (\forall v) \neg \delta(v, a) \), and hence \( \mathcal{B} \models \neg \delta(b, a) \), which is a contradiction.

For a model complete theory, all embeddings between models are elementary, so we have the following consequence.

4.8 COROLLARY. If a theory is model complete then it has AP.

In Section 3.2 we introduced the idea of moving from a theory \( T \) to a family of enriched theories \( T[\mathfrak{A}] \) for various structures \( \mathfrak{A} \). Model completeness is concerned with these enrichments for \( \mathfrak{A} \in Md(T) \). We now go to the other extreme, and consider what happens when we adjoin the diagram of submodels of \( T \).

4.9 DEFINITION. A theory \( T \) is submodel complete if the enriched theory \( T[\mathfrak{A}] \) is complete for each \( \mathfrak{A} \in \mathcal{S}(T) \).
Trivially, each submodel complete theory is model complete, but it should have further properties. The following result explains everything.

4.10 THEOREM. For each theory \( T \), the following are equivalent.

(i) \( T \) is submodel complete.

(ii) \( T \) is model complete and (the theory axiomatized by) \( T \cap \forall_1 \) has AP.

(iii) We have
\[(A, a) \equiv_0 (B, b) \implies (A, a) \equiv (B, b)\]
for all models \( A, B \) of \( T \) and points \( a \) from \( A \) and \( b \) from \( B \).

(iv) We have
\[(A, a) \equiv_0 (B, b) \implies (A, a) \equiv_1 (B, b)\]
for all models \( A, B \) of \( T \) and points \( a \) from \( A \) and \( b \) from \( B \).

(v) For each \( \forall_1 \)-formula \( \phi \) we have
\[T \vdash \phi \leftrightarrow \delta\]
for some quantifier-free formula \( \delta \).

(vi) \( T \) has EQ.

Proof. (i)\( \implies \) (ii). This is straight forward.

(ii)\( \implies \) (iii). Assuming (ii), consider any situation
\[(A, a) \equiv_0 (B, b)\]
where \( A, B \in Md(T) \). The point \( a \) of \( A \) need not enumerate a substructure of \( A \), but it certainly generates a substructure. (We merely take \( a \) and close off under the distinguished operations of \( A \).) In the same way, \( b \) generates a substructure of \( B \). The given relationship between \( a \) and \( b \) ensures that these two substructures are isomorphic. Let \( C \) be an isomorphic copy of these isomorphic substructure. Thus we have a wedge

\[
\begin{array}{c}
\mathcal{D} \\
A \downarrow \nwarrow \\
B & \equiv & C
\end{array}
\]

of embeddings, as to the left, where \( C \in \mathcal{S}(T) \) but \( A, B \in Md(T) \). Since \( T \cap \forall_1 \) has AP, this gives us a commuting square of embeddings, as to the right, where \( D \in \mathcal{S}(T) \). By a suitable enlargement we may suppose \( D \models T \). Thus, since \( T \) is model complete, the two right hand embeddings are elementary.

Notice how the two points \( a \) of \( A \) and \( b \) of \( B \) are situated in this square. By construction, there is a single point \( c \) of \( C \) which is sent to \( a \) and \( b \) by the left hand embeddings. But the square commutes, so there is a single point \( d \) of \( D \) to which \( a \) and \( b \) are sent by the right hand embeddings. Thus
\[(A, a) \equiv (D, d) \equiv (B, b)\]
to verify (iii).

(iii)⇒(iv). This is trivial.

(iv)⇒(v). Consider any $\forall_1$-formula $\phi$, and let $\Delta$ be the set of quantifier-free formulas $\delta$ with $\partial \delta \subseteq \partial \phi$ and such that $T \vdash \phi \rightarrow \delta$. Since $\Delta$ is close under conjunction, it suffices to show that

$$T \cup \Delta \vdash \phi$$

holds. Of course, to do this we need to assume (iv).

Consider any model $A \models T$, and consider any point $a$ of $A$ with $A \models \Delta(a)$. We must show that $A \models \phi(a)$ holds. To this end consider

$$T \cup \text{Diag}(A, a) \cup \{\phi(a)\}$$

where, of course, this diagram need not be the full diagram of $A$ (since $a$ need not enumerate the whole of $A$). By a simple argument, we see that this set is consistent. Thus we obtain a model $B$ of $T$ and a point $b$ of $B$ such that

$$(A, a) \equiv_0 (B, b) \quad B \models \phi(b)$$

hold. (We can’t say that $A \subseteq B$ since we are not working with the full diagram of $A$.) By (iv) we have

$$(A, a) \equiv_1 (B, b) \quad B \models \phi(b)$$

and hence $A \models \phi(a)$, as required.

(v)⇒(vi). This follows by induction on the complexity of $\phi$.

(vi)⇒(i). This is straightforward.

Exercises

4.10 Show that if a theory $T$ is $\forall_1$-axiomatizable and model complete then it has $EQ$. 
5

The back and forth technique

We are taking an historical line to set up the basics of Model Theory. In particular, we are investigating the question

What are the general logical principles behind quantifier elimination?

and in Theorem 4.10 we have come up with what looks like a reasonable answer. In particular, that answer indicates one reason why Model Theory perhaps should be developed a bit further.

Almost the only technique we have used so far is the compactness theorem, often in the form of a diagram argument. In this chapter we look at another important technique, the back-and-forth technique. This is important in a wider context, but here we concentrate on its model theoretic ramifications.

To help explain its model theoretic impact we set up a programme of analysing certain theories, and then carry out this programme for a few example theories.

We start with a theory $T$ which, as we will see, must have certain properties. In particular, it should be $\forall_2$-axiomatizable. (For almost all the examples we look at it is $\forall_1$-axiomatizable.) By analysing the properties of certain models of $T$ we produce a theory $T^*$ with the following properties.

(0) The theory $T^*$ is $\aleph_0$-categorical.
(1) The theory $T^*$ is complete.
(2) The theory $T^*$ is model complete.
(3) The theory $T^*$ has $\mathit{EQ}$.
(4) The theory $T^*$ is a ‘canonical’ companion of the theory $T$.

In more detail (4) says that $T^*$ has properties (0, 1, 2, 3) with $T \subseteq T^*$ and every model of $T$ is embeddable in a model of $T^*$. In Chapter 6 we investigate this canonical property in more detail. For now we simply take it at its face value.

Of course, to obtain such results the base theory $T$ must have rather special properties. Thus we choose our examples with some care.

Recall that a theory is $\aleph_0$-categorical if it has a countably infinite model and any two countable infinite models are isomorphic. Thus to verify property (0) we need a method by which we can show that two countable structures are isomorphic. That is precisely what the back-and-forth technique does. In fact, it does much more, as will be revealed as this chapter develops.

In Section 5.1 we set down the general properties of the back-and-forth technique. Then in Sections 5.2–5.4 we look at three examples which fit into the general programme outlined above.
As we will see, once we have verified property (0) the other four properties (1 – 4) follow by rather simple model theoretic arguments. Because of this and because of the nature of the examples chosen, it may look as though there is a tight connection between property (0) and properties (1 – 4). To dispel this idea in Section 5.5 we look at a theory which has properties (1, 2, 3), but not (0), and yet the back-and-forth technique is the crucial tool that helps us analyse the theory.

5.1 The technique

In this section we describe the generalities of the back and forth technique. You may find this a bit abstract. If so you should begin to read Sections 5.2 and 5.3 in parallel with this. Those two sections (and the following sections) describe particular examples of the technique. They should help you understand these generalities.

The technique is used to produce isomorphisms bit by bit.

5.1 Definition. Let $\mathfrak{A}$, $\mathfrak{B}$ be structures for some language $L$. A partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ is a bijection $f : U \rightarrow V$

from a subset $U$ of $\mathfrak{A}$ to a subset $V$ of $\mathfrak{B}$ which is an isomorphism ‘as far as it goes’.

Of course, this attempted definition more or less begs the question. What does the phrase ‘as far as it goes’ mean. We could begin to spell it out along the lines of Definition 3.4 but that is rather tedious (and almost certainly will be incorrect). Fortunately we now know a better way to make the notion precise.

Consider the situation of Definition 5.1. Let $\mathfrak{a}$ be an enumeration of $U$ and let $\mathfrak{b}$ be the corresponding enumeration of $V$. In more detail, if for some index set we have $a = (a_i \mid i \in I)$ for the enumeration of $U$, then $b = (f(a_i) \mid i \in I)$ is the corresponding enumeration of $V$. We can now enrich the parent language $L$ by a family of constants $(K_i \mid i \in I)$ indexed by $I$. Of course $(\mathfrak{A}, a)$ and $(\mathfrak{B}, b)$ are structures for the enriched language.

With this we can be more precise.

5.2 Definition. Let $\mathfrak{A}$, $\mathfrak{B}$ be structures for some language $L$. A partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ is a pair $(a, b)$ of enumerations (a from $\mathfrak{A}$ and b from $\mathfrak{B}$) such that $(\mathfrak{A}, a) \equiv_0 (\mathfrak{B}, b)$ holds.
Think of \( a \) and \( b \) as listing a part \( U \) of \( \mathfrak{A} \) and a part \( V \) of \( \mathfrak{B} \), respectively. Then the relation 

\[(\mathfrak{A}, a) \equiv_0 (\mathfrak{B}, b)\]

certainly sets up a bijection between \( U \) and \( V \). Furthermore, this bijection is an isomorphism ‘as far as it goes’. Notice that we do not assume that \( a \) and \( b \) enumerate substructures of \( \mathfrak{A} \) and \( \mathfrak{B} \). However, the two will generate substructures, and these substructures will be canonically isomorphic.

The general idea is to take such a partial isomorphism and extend it a bit further. By repeatedly extending we hope to produce a full isomorphism.

In the following definition we revert to the style of Definition 5.1.

5.3 **DEFINITION.** Let \( \mathfrak{A}, \mathfrak{B} \) be structures for some language \( L \).

(a) A **back-and-forth system** for the pair \( \mathfrak{A}, \mathfrak{B} \) is a non-empty set \( P \) of partial isomorphisms with the following two properties.

(back) For each member of \( P \)

\[f : U \longrightarrow V\]

and each element \( y \) of \( \mathfrak{B} \), there is an element \( x \) of \( \mathfrak{A} \) together with an extension

\[f^+ : U \cup \{x\} \longrightarrow V \cup \{y\}\]

of \( f \) in \( P \).

(forth) For each member of \( P \)

\[f : U \longrightarrow V\]

and each element \( x \) of \( \mathfrak{A} \), there is an element \( y \) of \( \mathfrak{B} \) together with an extension

\[f^+ : U \cup \{x\} \longrightarrow V \cup \{y\}\]

of \( f \) in \( P \).

We often abbreviate ‘back-and-forth’ to ‘b&f’.

(b) We write

\[\mathfrak{A} \equiv_p \mathfrak{B}\]

and say \( \mathfrak{A} \) and \( \mathfrak{B} \) are **potentially isomorphic** if there is at least one back-and-forth system for the pair \( \mathfrak{A}, \mathfrak{B} \).

It’s time for a couple of simple examples.

5.4 **EXAMPLES.** Let \( \mathfrak{A}, \mathfrak{B} \) be a pair of structures and let \( f \) be a full isomorphism from \( \mathfrak{A} \) to \( \mathfrak{B} \).

(a) The single function \( \{f\} \) is a back-and-forth system for \( \mathfrak{A}, \mathfrak{B} \).

(b) The set of all bijections 

\[f_U : U \longrightarrow f[U]\]

for finite subsets \( U \) of \( \mathfrak{A} \) is a back-and-forth system for \( \mathfrak{A}, \mathfrak{B} \).
Notice that potential isomorphisms does not require that the two structures have the same cardinality.

The relation \( \equiv_p \) should be compared with the semantic relations
\[
\equiv_n \quad \equiv_\omega
\]
(for \( n < \omega \)) and the isomorphism relation
\[
\equiv_\infty
\]
defined in Section 3.1. We have used a similar notation for these relations because of the following result.

5.5 THEOREM. For each pair \( \mathfrak{A}, \mathfrak{B} \) of structures (for some language), both the implications
\[
\mathfrak{A} \equiv_\infty \mathfrak{B} \implies \mathfrak{A} \equiv_p \mathfrak{B} \implies \mathfrak{A} \equiv_\omega \mathfrak{B}
\]
hold.

Proof. As in Example 5.4(a), the first implication is trivial.

For the second implication suppose we have a back-and-forth system \( P \) for \( \mathfrak{A}, \mathfrak{B} \). We show
\[
[n] \quad (\mathfrak{A}, a) \equiv_n (\mathfrak{B}, b)
\]
for each \( n < \omega \) and each \( (a, b) \in P \). We proceed by induction on \( n \).

The base case, \([0]\), is trivial by the defining property of \( P \).

For the induction step, \([n] \Rightarrow [n + 1] \), suppose consider any \( \forall_n \)-formula
\[
\phi(v_1, \ldots, v_r, w_1, \ldots, w_s)
\]
with
\[
\mathfrak{A} \models (\exists v_1, \ldots, v_r)\phi(v_1, \ldots, v_r, a_1, \ldots, a_s)
\]
where the two points
\[
(a_1, \ldots, a_s) \text{ of } \mathfrak{A} \quad (b_1, \ldots, b_s) \text{ of } \mathfrak{B}
\]
are matched by some member of \( P \). In other words, there is some member
\[
f : U \longrightarrow V
\]
of \( P \) with
\[
a_1, \ldots, a_s \in U \quad b_1, \ldots, b_s \in V
\]
and
\[
f(a_j) = b_j
\]
for each \( 1 \leq j \leq s \). We show that
\[
\mathfrak{B} \models (\exists v_1, \ldots, v_r)\phi(v_1, \ldots, v_r, b_1, \ldots, b_s)
\]
holds.

There are elements \( x_1, \ldots, x_r \) of \( \mathfrak{A} \) such that
\[
\mathfrak{A} \models \phi(x_r, \ldots, x_1, a_1, \ldots, a_s)
\]
5.1. The technique

holds. By repeated use of the ‘forth’ property there are elements \( y_1, \ldots, y_r \) of \( \mathcal{B} \) together with some extension

\[
f^+ : U^+ \longrightarrow V^+
\]

of \( f \) in \( P \) with

\[
x_1, \ldots, x_r \in U^+ \quad y_1, \ldots, y_r \in V^+
\]

and

\[
f(x_i) = y_i
\]

for each \( 1 \leq i \leq r \).

We have

\[
(\mathfrak{A}, x_1, \ldots, x_r, a_1, \ldots, a_s) \equiv_n (\mathfrak{B}, y_1, \ldots, y_r, b_1, \ldots, b_s)
\]

by the induction hypothesis \([n]\). This gives

\[
\mathfrak{B} \models \phi(y_1, \ldots, y_r, b_1, \ldots, b_s)
\]

and hence

\[
\mathfrak{B} \models (\exists v_1, \ldots, v_r)\phi(v_1, \ldots, v_r, b_1, \ldots, b_s)
\]

as required.

This, with a symmetric argument, gives \([n + 1]\). ■

The relation \( \equiv_p \) sits between the two extremes \( \equiv_\infty \) and \( \equiv_\omega \), and in appropriate circumstances it can agree with either of these extremes.

Originally, back-and-forth systems were invented to generate full isomorphisms. Here is that first use.

5.6 THEOREM. We have

\[
\mathfrak{A} \equiv_p \mathfrak{B} \implies \mathfrak{A} \equiv_\infty \mathfrak{B}
\]

for all countable structures \( \mathfrak{A}, \mathfrak{B} \).

Proof. Let \( P \) be a back-and-forth system for the two countable structures \( \mathfrak{A} \) and \( \mathfrak{B} \). We use \( P \) to generate a full isomorphisms \( f \) from \( \mathfrak{A} \) to \( \mathfrak{B} \).

Let \( A \) be the carrier of \( \mathfrak{A} \) and let \( B \) be the carrier of \( \mathfrak{B} \). Since both \( A \) and \( B \) are countable there are enumerations

\[
(x_i \mid i < \omega) \quad (y_i \mid i < \omega)
\]

of these sets. (These enumerations can be quite arbitrary.) We produce an ascending chain

\[
f_0 \subseteq f_1 \subseteq \cdots \subseteq f_i \subseteq \quad i < \omega
\]

of partial isomorphism

\[
f_i : U_i \longrightarrow V_i
\]

in \( P \) and we arrange that

\[
x_i \in U_{i+1} \quad y_i \in V_{i+1}
\]

for each \( i < \omega \). The union

\[
f = \bigcup\{f_i \mid i < \omega\}
\]
is then the required isomorphism.

Let $f_0$ be any member of $P$.

Suppose we have produced $f_i$. By first going forth (to deal with $x_i$) and then coming back (to deal with $y_i$) we obtain $f_{i+1}$. ■

As we will see this back-and-forth technique is quite powerful. For instance, here is how we get properties (0, 1).

5.7 THEOREM. Let $T$ be a theory in a countable language and suppose $T$ has no finite models. Suppose also that

$$A \equiv_p B$$

for all countable models $A, B$ of $T$. Then $T$ is $\aleph_0$-categorical and complete.

You should be able to see the proof of this immediately. The back-and-forth technique can also deal with model completeness, but that is not so obvious.

Exercises

5.1 Show that potential isomorphism, $\equiv_p$, is an equivalence on the class of all structures (for a given language).

5.2 Show that for each b&f system $P$ between a pair $A, B$ of structures, we have

$$(A, a) \equiv (B, b)$$

for each $(a, b) \in P$.

5.3 Prove Theorem 5.7.

5.2 Linear orders

In Section 2.2 we considered the theory of dense linear orders without end points, which for brevity we referred to as lines. We exhibited, or at least showed the existence of, an effective algorithm for eliminating quantifiers within that theory. In this section we prove a model theoretic version of that result, and obtain quite a bit more information about the theory. In particular, we show that the theory has $EQ$ but we don’t produce an effective algorithm. In the long run, what use is that algorithm?

We work within a language which has just one signature symbol, a 2-placed relation symbol $\leq$ which, as usual, we write as a infix. We look at two theories formalized within that language.

5.8 DEFINITION. Let $T(Lord)$ and $T(Line)$ be, respectively, the theory of linear orders and the theory of dense linear orders without end points. ■

As indicated we continue to use ‘line’ as an abbreviation for ‘dense linear order without end points’.

We investigate the interaction between $T(Lord)$ and $T(Line)$ and produce some information about the spectrum of models of these theories.

Observe that

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$$

are models of $T(Lord)$ and the last two are models of $T(Line)$.  

5.9 **THEOREM.** The theory $T(\text{Lord})$ is $\forall_1$-axiomatizable, and the theory $T(\text{Line})$ is $\forall_2$-axiomatizable.

The proof of this is trivial. We just write down the axioms.

Next we develop just a little bit of the arithmetic of linear order types. This is important in its own right and at some stage you should learn more about the topic. Here we stick to the basics.

Each linear ordered set $A$ has an isomorphism type $\alpha$. This type $\alpha$ doesn’t tell us what $A$ is but does tell what $A$ looks like ‘up to isomorphism’. For instance, we often use $\omega$ for the isomorphism type of $\mathbb{N}$ (as a linear ordering).

Given two linear ordered sets $A$ and $B$ with types $\alpha$ and $\beta$, there several constructions which convert the pair $A, B$ into another linear ordered set where the resulting order type depends only on $\alpha, \beta$ and not the actual inner details of $A, B$. Here are the two simplest such constructions.

5.10 **DEFINITION.** (+) For order types $\alpha, \beta$ let

$$\beta + \alpha$$

be the order type obtained by taking disjoint copies $A, B$ of $\alpha, \beta$ and sticking these end to end

$$B \cup A$$

without overlap. This gives the sum of the two order types.

(·) For order types $\alpha, \beta$ let

$$\beta \cdot \alpha$$

be the order type obtained by taking a copy $A$ of $\alpha$ and replacing each of its elements by a (different) copy of $B$ without overlap. This gives the product of the two order types. ■

The sum of two order types is easy to understand, but the product needs a little bit more explanation.

To view $\beta \cdot \alpha$ first take a copy of $\alpha$.

\[ \bullet \ \ldots \ \bullet \ \ldots \ \bullet \ \ldots \ \bullet \ \ldots \ \bullet \]

Here I have drawn the linear order as though it may have gaps. That is merely to help me draw the next picture. Now take many disjoint copies of $\beta$. Let me use

\[ \sim \ast \sim \]

as a typical example of such a linear order. Now replace each element of $\alpha$ by a copy of $\beta$. More prosaically, think of $\alpha$ as a washing line and hang a $\sim \ast \sim$ from each element.

\[ \bullet \ \ldots \ \bullet \ \ldots \ \bullet \ \ldots \ \bullet \]

\[ \sim \ast \sim \quad \sim \ast \sim \quad \sim \ast \sim \quad \sim \ast \sim \quad \sim \ast \sim \]

These dangling disjoint copies of $\sim \ast \sim$ form a linear ordering of type $\beta \cdot \alpha$. Within each copy the elements are ordered as in the copy. Two elements in different copies are
ordered by the two elements in \( \alpha \) from which these dangle. (Of course, for the pictorial representation you must assume that there is no more than a light breeze.)

There is a whole arithmetic of linear order types. Exercise 5.6 gives some of the basic rules.

Before we leave this let’s observe that

\[
\mathbb{N} \quad \mathbb{Z} \quad \mathbb{Q}
\]

are examples of three important order types. What do you make of the following.

\[
\mathbb{N} + \mathbb{Z} \cdot \mathbb{Q}
\]

You will meet this again later.

These ideas give us a couple of structural properties of \( T(\text{Lord}) \).

5.11 THEOREM. The theory \( T(\text{Lord}) \) has JEP and AP.

Proof. (jep) This is easy. Given two linear orders \( A, B \), the disjoint sum \( B + A \) (with \( B \) to the left and \( A \) to the right) embeds both \( A \) and \( B \).

(ap) Consider a wedge of embeddings

\[
\begin{array}{c}
B \\
A \\
C
\end{array}
\]

between linear ordered sets. By replacing \( B, C \) by isomorphic copies we may assume that

\[
B \cap C = A
\]

and that the two embedding are insertions. It now suffices to merge \( C - A \) into \( B \) in a way that preserves the order of \( C \) and its relationship with \( A \). To do that we build a larger and larger linearly ordered set which includes \( B \) and eventually includes the whole of \( C \). We build this larger linearly ordered set element by element of \( C - A \).

We enumerate \( C - A \)

\[
(c_i \mid i < \gamma)
\]

using a suitable ordinal \( \gamma \) and then proceed by recursion over \( i \). Each step is the same.

Suppose we have to deal with some element \( c \in C \). It could be the some earlier elements have been merged into \( B \), and of course \( A \) is there already. Compare \( c \) with the elements of \( C \) already merged into \( B \). This gives us two disjoint sets of elements of \( B \); the set \( L \) of those elements that correspond to some element of \( C \) to the left of \( c \), and the set \( R \) of those elements that correspond to some element of \( C \) to the right of \( c \). Of course, there will be some element on \( B \) that don’t occur in \( L \cup R \). We now slip \( c \) anywhere between \( L \) and \( R \). We don’t look for a suitable element of \( b \), we merely insert a new element. There may be some elements of \( B \) already in this gap, but they don’t interfere with this construction.

This proof uses a rather crude version of a method we employ in the b&f technique. The proof of the next result is a better example of the method. As with most of results of this kind its proof can be split into two parts; a preliminary 1-step construction, and then an iteration of that construction.
5.12 LEMMA. Let $A$ any linearly ordered set and let $B$ be any line. Let $X$ be a finite subset of $A$ viewed as a linearly ordered set, and let

$$X \xrightarrow{f} B$$

be an embedding of $X$ into $B$.

For each element $a \in A$ there is an embedding

$$X \cup \{a\} \xrightarrow{f'} B$$

that extends $f$.

Proof. Recall that since $f$ is an embedding we have

$$x \leq y \iff f(x) \leq f(y)$$

for each $x, y \in X$.

We may suppose that $a \notin X$ (for otherwise we let $f' = f$).

The element $a \in A$ splits the set $X$ into two disjoint parts $L, R$. The set $L$ consists of those $x \in X$ with $x < a$, and the set $R$ consists of those $x \in X$ with $a < x$. Both these sets are finite and there are three possibilities

(left) $L = \emptyset$  (gap) $L \neq \emptyset \neq R$  (right) $R = \emptyset$

and we deal with these cases in slightly different ways, but in each case we look at the image

$$f[L] \quad f[R]$$

of the two sets in $B$.

Consider the case (gap). Let $l$ be the maximum of $L$ and let $r$ be the minimum of $R$. We have

$$l < a < r$$

and also

$$f(l) < f(r)$$

in $B$. All we require is some element $b$ of $B$ which is strictly between $f(l)$ and $f(r)$. Since $B$ is dense, there is such a element $b$.

Consider the case (left). Then $R = X$ with $a < x$ for each $x \in X$. Let $r$ be the minimum of $R$. We have

$$a < r$$

and we require some element $b$ of $B$ with $b < f(r)$. Since $B$ has no first element, there is such a element $b$.

The case (right) is dealt with in the same way.  ■

By iterating this trick we obtain the following.

5.13 THEOREM. Each countable linearly ordered set is embeddable in each line.
Proof. Let $A$ a countable linearly ordered set and let $B$ be any line. We produce an embedding

$$A \longrightarrow B$$

bit by bit.

Since $A$ is countable is has an enumeration

$$(a_r | r < \omega)$$

where, of course, this has no connection with the given comparison on $A$. For each $r < \omega$ let

$$X(r) = \{a(0), \ldots, a(r)\}$$

to give an ascending chain

$$X(0) \subseteq X(1) \subseteq \cdots \subseteq X(r) \subseteq \cdots \quad (r < \omega)$$

of finite subsets of $A$ where the union is $A$. Each step $X(r) \to X(r + 1)$ is given by one extra element (namely, $a(r + 1)$).

Choose any element of $B$, and call it $f(a(0))$. This gives us an embedding

$$X(0) \overset{f_0}{\longrightarrow} B$$

from which we can start our construction.

By repeated use of Lemma 5.12 we build an ascending chain

$$
\begin{array}{c}
\vdots \\
X(r) \overset{f_r}{\longrightarrow} B \\
\vdots \\
\vdots \\
X(2) \overset{f_2}{\longrightarrow} B \\
\vdots \\
x(1) \overset{f_1}{\longrightarrow} B \\
X(0) \overset{f_0}{\longrightarrow} B
\end{array}
$$

of embeddings where the left hand chain eventually exhausts $A$. The union of this chain is an embedding of $A$ into $B$. □

The countability in this result is not a serious restriction

5.14 Corollary. Each linearly ordered set $A$ is embeddable in a line.
5.2. Linear orders

Proof. Consider any linearly ordered set $A$. Let $a$ be an enumeration of the whole of $A$. By Theorem 5.13 (or even a finite iterated use of Lemma 5.12), we see that

$$T(\text{Line}) \cup \text{Diag}(A, a)$$

is finitely satisfiable, and hence has a model, to give the required result. ■

Now that we are getting used to this tune, we can start to swing it.

5.15 THEOREM. For each pair $A, B$ of lines, the family of all finite partial isomorphisms from $A$ to $B$ is a b&f system.

Proof. Consider any finite partial isomorphism

$$X \xrightarrow{f} Y$$

from $A$ to $B$. A use of Lemma 5.12 enables us to go forth, and then a second use (with the roles reversed) enables us to come back. ■

This result has several almost immediate consequences.

5.16 COROLLARY. We have

$$A \equiv_p B$$

for each pair $A, B$ of lines.

And a particular case of this deserves a higher status.

5.17 THEOREM. The theory $T(\text{Line})$ is $\aleph_0$-categorical.

And so does another simple consequence.

5.18 THEOREM. The theory $T(\text{Line})$ is complete.

A variant of this result gives the following.

5.19 LEMMA. Let $A, B$ be a pair of lines. Then we have

$$(A, a) \equiv_0 (B, b) \implies (A, a) \equiv_p (B, b)$$

for each pair $(a, b)$ finite sequences from $A$ and $B$.

Proof. Suppose

$$(A, a) \equiv_0 (B, b)$$

and let $X \subseteq A$ and $Y \subseteq B$ be these finite sequences viewed as linearly ordered sets. Thus we have a partial isomorphism

$$X \xrightarrow{f} Y$$

from $A$ to $B$. By Theorem 5.15 this is the base of a b&f system between the two structures $(A, a)$ and $(B, b)$. Thus

$$(A, a) \equiv_p (B, b)$$

as required. ■

A particular case gives the following.
5.20 **THEOREM.** The theory $T(\text{Line})$ is model complete.

**Proof.** Consider any pair of lines $A \subseteq B$ one included in the other. Consider any finite list $a$ of elements of $A$. We have $(A, a) \subseteq (B, b)$ so that $(A, a) \equiv_0 (B, b)$ and hence $(A, a) \equiv_p (B, b)$ by Lemma 5.19 with $(A, a) \equiv (B, b)$ as a special case. By letting $a$ vary through all possible finite sequences we have $A \prec B$ as required. ■

Using the structural properties given by Theorem 5.11 we can strengthen Theorem 5.19.

5.21 **THEOREM.** The theory $T(\text{Line})$ has $\text{EQ}$.

**Proof.** By Theorems 4.10 and 5.19 it suffices to show that the theory (axiomatized by)

$$T(\text{Line}) \cap \forall_1$$

has $\text{AP}$. By Corollary 5.14 this theory is just $T(\text{Lord})$, and by Theorem 5.11 this theory has $\text{AP}$. ■

With this we have more or less proved the following.

5.22 **THEOREM.** The theory $T(\text{Line})$ is a ‘canonical’ companion of $T(\text{Lord})$.

**Proof.** We have $T(\text{Lord}) \subseteq T(\text{Line})$ where $T(\text{Line})$ has properties $(0, 1, 2, 3)$. Thus it suffices to show that each model of $T(\text{Lord})$ is embeddable in some model of $T(\text{Line})$. This is just Corollary 5.14. ■

You should observe the global structure of this section. Once we have the b&f property almost everything else follows very quickly.
Exercises

5.4 Make sure you can write down the axioms needed to prove Theorem 5.9.

5.5 Consider the product $\beta \cdot \alpha$ of two order types. Let $A, B$ be linear orderings of types $\alpha, \beta$, respectively. The product can be represented by a certain linear ordering of the cartesian product $B \times A$.

Write down this linear ordering.

5.6 Consider the arithmetic of order types.

(a) Show that addition is associative but not commutative.

(b) Show that multiplication is associative but not commutative.

(c) Investigate the distributive laws.

5.3 Graphs

In this section we look at another pair of theories

$$T(Gph) \subseteq T(Rdm)$$

which have the relationship as described in the general programme set out in the preamble to this chapter. The broad development follows quite closely that of Section 5.2 but, of course, with different structures involved.

We are concerned with graphs. There are various different notions of graph around, so the first thing we should do is make sure we know which one we use here.

5.23 DEFINITION. In this context a graph is a structure

$$(A, E)$$

where $A$ is a (non-empty) set and $E$ is a binary relation on $A$ which is both symmetric and irreflexive.

Let $T(Gph)$ be the theory of all such graphs. ■

The signature of the underlying language has just one symbol, a 2-placed relation symbol which, as is customary, we write as an infix. A structure for this language is a graph precisely when it has the following two properties.

(Sym) $$(\forall x, y)[xEy \rightarrow yEx]$$

(Irr) $$(\forall x)[\neg(xEy)]$$

Since these are sentences of the underlying language we immediately have the following rather trivial observation.

5.24 THEOREM. The theory $T(Gph)$ of graphs is $\forall_1$-axiomatizable.

Given a graph $(A, E)$ we may picture it as a collection of nodes, the members of $A$, some of which are connected to indicate the relation $E$. Informally we may think of two connected nodes as mates. Given two nodes $x, y$ we can write

$$x \rightarrow y \quad x \leftarrow y$$
5. The back and forth technique

to show that \( x \) and \( y \) are mates are not mates respectively. Throughout the analysis we will use this notation (and a couple of variants) rather than the more formal \( E \).

You should not confuse this notion of graph that of a directed graph. That too consists of a collection of nodes, but these are connected by directed edges. Here we have undirected connections.

Our aim is to locate a ‘canonical’ companion of \( T(Gph) \) and obtain some information about that theory. To do that we use the two standard structural properties.

5.25 THEOREM. The Theory \( T(Gph) \) has both JEP and AP.

Proof. To deal with AP consider any wedge of embeddings between graphs.

\[
\begin{array}{c}
B \\
f \\
A \\
g \\
C
\end{array}
\]

By taking appropriate isomorphic copies we may suppose that \( f \) and \( g \) are inclusions and \( B \cap C = A \) holds. We now furnish the set

\[
D = (B - A) \cup A \cup (C - A)
\]

in the obvious way to produce a graph and close the wedge.

Dealing with JEP is even easier. We merely take a disjoint union of two graphs. ■

Where might we find a ‘canonical’ companion of \( T(Gph) \), if such a theory exists. We look for graphs that are sufficiently closed. By that we mean that if it is possible to extend the graph to include a certain configuration, then that configuration already exists. Here is the precise definition of that informal notion.

5.26 DEFINITION. A graph is random is for each disjoint pair \( L, R \) of finite sets of nodes we have

\[
(\forall x \in L)[x \leftrightarrow z] \quad (\forall y \in R)[z \rightarrow y]
\]

for at least one node \( z \).

Let \( T(Rdm) \) be the theory of all random graphs. ■

We sometimes write

\[
L \rightarrow z \rightarrow R
\]

to indicate that the node \( z \) satisfies the condition for the particular sets \( L \) and \( R \).

Does \( T(Rdm) \) have a sensible axiomatization? Look again at the definition. If we fix the size of the two finite sets \( L \) and \( R \) then it is fairly easy to write down the condition

\[
L \rightarrow z \rightarrow R
\]

as a single \( \exists_1 \)-formula. By univelsally quantifying out \( L \) and \( R \), for that size of sets, we obtain a single \( \forall_2 \)-sentence. By taking those sentences for all sizes of sets we obtain the following.
5.27 **THEOREM.** The theory $T(Rdm)$ of random graphs is $\forall_2$-axiomatizable.

Trivially we have $T(Gph) \subseteq T(Rdm)$. Our aim is to show that $T(Rdm)$ has a particularly important position in the family of all extensions of $T(Gph)$.

Just why this kind of graph is called random will have to remain a mystery for a while. It is beyond our remit to worry about such things.

However, this definition does pose two immediate questions we should have a little worry about. Are there any random graphs? In fact some basic model theory give a rather easy proof of existence, but as yet we don’t have that machinery. We deal with that in Chapter 6.

Here we exhibit a particular random graph. We find this in an unexpected place, the lower levels of the Zermelo hierarchy. To produce this we take a slight detour.

5.28 **DEFINITION.** Let

$$V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_r \subseteq \cdots \quad (r < \omega)$$

be generated by

$$V_0 = \emptyset \quad V_{r+1} = \mathcal{P}V_r$$

for each $r < \omega$. Let

$$V_\omega = \bigcup\{V_r \mid r < \omega\}$$

to obtain the family of hereditarily finite sets.

It is easy to check that

$$V_r \subseteq V_{r+1}$$

and hence $V_\omega$ is the union of an ascending chain of finite sets, as indicated. Observe that each member $x \in V_\omega$ is a finite set, and each member of $x$ is a finite sets which is also in $V_\omega$. Thus each $x \in V_\omega$ can be represented as a finite rooted tree. In fact, $V_\omega$ is just the collection of all possible ‘shapes’ a finite rooted tree can take. (To make this more precise we have to be careful when two ‘shapes’, are the same. Fortunately, we don’t need to do that here.)

We convert $V_\omega$ into a graph in what may seem a strange way.

5.29 **DEFINITION.** Let $\sim$ be the relation on $V_\omega$ given by

$$x \sim y \iff x \in y \text{ or } y \in x$$

for $x, y \in V_\omega$.

By construction this relation $\sim$ is symmetric. It is also irreflexive since there is no $x \in V_\omega$ with $x \in x$. Thus we do have a graph.

5.30 **THEOREM.** The graph $(V_\omega, \sim)$ is random.
Proof. Observe that any finite subset of $V_\omega$ is also a member of $V_\omega$.
Consider a disjoint pair $L, R$ of finite subsets of $V_\omega$. We require
\[ L \sim a \sim R \]
for some $a \in V_\omega$.
Consider the family $A$ of all those $a \in V_\omega$ with $L \subseteq a$. This family $A$ is infinite. For each $a \in A$ we have
\[ x \in L \implies x \in a \implies x \sim a \]
(for $x \in V_\omega$), and so it suffices to show that
\[ a \sim R \]
for at least one $a \in A$.
For an arbitrary $a \in A$ suppose that the extra requirement doesn’t hold. Thus
\[ y \in a \text{ or } a \in y \]
for some $y \in R$. We show that this obstructs only finitely many members of $A$. Since $A$ is infinite we have more than enough $a \in A$ left to satisfy the requirement.
The set $V_\omega$ is infinite and the set $R$ is finite. Thus the difference
\[ B = V_\omega - R \]
is infinite. Since $L$ and $R$ are disjoint we have $L \subseteq B$.
Consider any
\[ a = L \cup b \]
where $b$ is a finite subset of $B$. Trivially we have $a \in A$, and $a$ is a finite subset of $B$.
Since this $a$ is a finite subset of $V_\omega$, it is also a member of $V_\omega$. For each $y \in V_\omega$ we have
\[ y \in a \implies y \in b \implies y \notin R \]
and hence such a set $a$ avoids the left hand obstruction. Since $B$ is infinite this still leaves us with infinitely many possible choices for $a$ to try to avoid the right hand obstruction.
For such an $a$ suppose $a \in y$ for some $y \in R$. Since $R$ is finite and each $y \in R$ is finite, this obstruction places $a$ as a member of some finite set. In other words, the right hand obstruction rules out only finitely many possible $a$, and again we still have plenty left from which to choose.

Since the set $V_\omega$ is countable infinite we have the following.

5.31 COROLLARY. The theory $T(Rdm)$ has a countably infinite model.

In due course we will show that each random graph has rather nice b&f properties. To prepare for that let’s first look at the forth properties.
In the next couple of results we consider an arbitrary random graph $(V, \sim)$. As here we write $\sim$ for its mating relationship. This graph need not be countable.
5.32 LEMMA. Let $V$ be any random graph. Let $A$ be any graph, let $B \subseteq A$ be any finite subgraph, and let

\[ B \xrightarrow{f} V \]

be any embedding. For each $a \in A$ there is some $c \in V$ such that $f$ extends to an embedding

\[ B \cup \{a\} \xrightarrow{f'} V \]

of the larger subgraph of $A$ by setting $f'(a) = c$.

Proof. Fix $a \in A$. Consider the range $f[B]$ of $f$ in $V$. This is finite (since $B$ is finite). We use $a$ to split $f[B]$ into two parts $L, R$ by

\[ f(b) \in L \iff b \|-- a \quad a \|-- b \iff f(b) \in R \]

for each $b \in B$. Here we are using the mating relation on $A$. Trivially, $L$ and $R$ are finite and disjoint.

Since $V$ is random there is some $c \in V$ such that

\[ b \|-- a \implies f(b) \in L \implies f(b) \sim c \quad a \|-- b \implies f(b) \in R \implies f(b) \not\sim c \]

for each $b \in B$. In fact, by the construction of $L$ and $R$, these implications are equivalences. Thus $c$ is a suitable element of $V$ by which we can extend the given embedding $f$. ■

With this trick, by repeatedly going forth we obtain the following.

5.33 THEOREM. Let $V$ be any random graph. Then each countable graph is embeddable in $V$.

Proof. Let $A$ be any countable graph (countable model of $T(Gph)$). We must produce an embedding of $A$ into $V$. Let

\[ (a(r) \mid r < \omega) \]

be any enumeration of $A$ (with repetitions allowed). For each $r < \omega$ let $A_r$ be the subgraph of $A$ on the set $\{a(0), \ldots, a(r)\}$. This gives us an ascending chain

\[ A(0) \subseteq A(1) \subseteq \cdots \subseteq A(r) \subseteq \cdots \quad (r < \omega) \]

of finite subgraphs of $A$ where the union is $A$. Each step $A_r \leftrightarrow A_{r+1}$ is given by one extra element (namely, $a(r+1)$).

Choose any element of $V$, and call it $f(a(0))$. This gives us an embedding

\[ A(0) \xrightarrow{f_0} V \]

from which we can start our construction.
By repeated use of Lemma 5.32 we build an ascending chain

\[ \begin{array}{cccc}
A(r) & f_r & V \\
A(2) & f_2 & V \\
A(1) & f_1 & V \\
A(0) & f_0 & V \\
\end{array} \]

of embeddings where the left hand chain eventually exhausts \( A \). The union of this chain is an embedding of \( A \) into \( V \).  

A simple diagram argument enables us to deal with all graphs

5.34 COROLLARY. Each graph is embeddable in at least one random graph.

Proof. For an arbitrary graph \( A \) look at

\[ T(Rdm) \cup Diag(A, a) \]

where \( a \) is an enumeration of \( A \). It suffices to show that this set is satisfiable.

Consider any finite part \( \Sigma \) of \( Diag(A, a) \). This set \( \Sigma \) is satisfied in some finite subgraph of \( A \) (the subgraph on the elements named in \( \Sigma \)). By Theorem 5.33 this subgraph is embeddable in at least one random graph, to show that

\[ T(Rdm) \cup \Sigma \]

is satisfiable. A standard use of compactness now gives the required result.

Since

\[ T(Gph) \subseteq T(Rdm) \]

this show \( T(Rdm) \) has a rather tight connection with \( T(Gph) \). Our aim is to show that \( T(Rdm) \) is a ‘canonical’ companion of \( T(Gph) \).

We now look at the b&f properties of random graphs. Here is the crucial one-step back and forth constuction.

5.35 LEMMA. Let \( U, V \) be any pair of random graphs. Let \( A \) be a finite subgraph of \( U \) and let \( B \) be a finite subgraph of \( V \). Let

\[ \begin{array}{cccc}
U & V \\
A & B \\
\end{array} \]
be any isomorphism between the two subgraphs, a partial isomorphism from $U$ to $V$. Let $a$ be any element of $U$ and let $b$ be any element of $V$. Then there are elements $a'$ of $U$ and $b'$ of $V$ such that setting

$$f^+(a) = b' \quad f^+(a') = b$$

produces an extension

$$U \quad V$$

$$\downarrow \quad \downarrow$$

$$A \cup \{a, a'\} \xrightarrow{f^+} B \cup \{b', b\}$$

$$\downarrow \quad \downarrow$$

$$A \xrightarrow{f} B$$

of the partial isomorphism $f$.

Proof. We produce the extension $f^+$ by two applications of Lemma 5.32. First of all we use the element $a \in U$ to go forth to produce an extension $f^-$ of $f$ with

$$f^-(a) = b'$$

for some element $b' \in V$.

The inverse of this partial isomorphism from $U$ to $V$ is a partial isomorphism from $V$ to $U$. Using the element $b \in V$ we come back to produce an extension $f^+$ of $f^-$ with

$$f^+(a') = b$$

for some element $a' \in U$. ■

In short, this is what we have just proved.

5.36 COROLLARY. For each pair $U, V$ of random graphs the set of all finite partial isomorphisms between the two is a b&f system.

This gives us the following.

5.37 THEOREM. For each pair $U, V$ of random graphs we have

$$U \equiv_p V$$

and, in particular, the theory $T(Rdm)$ is $\aleph_0$-categorical and complete.

As usual we can get a bit more out of this.

5.38 SCHOLIUM. Let $U, V$ be a pair of random graphs. Let $a$ list a finite part of $U$ and let $b$ list a finite part of $V$, and suppose these two finite graphs are isomorphic. Then

$$(U, a) \equiv_p (V, b)$$

holds.
Now consider any random graphs $U \subseteq V$. Let $a$ be any finite part of $U$. Then

$$(U, a) \equiv_p (V, a)$$

to give

$$(U, a) \equiv (V, a)$$

and hence $U \prec V$. Thus we have the following.

5.39 THEOREM. The theory $T(Rdm)$ is model complete.

In fact, we can strengthen Theorem 5.39.

5.40 THEOREM. The theory $T(Rdm)$ has EQ.

Proof. By Theorems 4.10 and 5.19 it suffices to show that the theory (axiomatized by)

$$T(Rdm) \cap \forall_1$$

has AP. By Corollary 5.36 his theory is just $T(Gph)$, and by Theorem 5.25 this theory has AP.  

With this we have more or less proved the following.

5.41 THEOREM. The theory $T(Rdm)$ is a ‘canonical’ companion of $T(Gph)$.

Proof. We have $T(Gph) \subseteq T(Rdm)$ where $T(Rdm)$ has the properties (0, 1, 2, 3). Thus it suffices to show that each model of $T(Gph)$ is embeddable in some model of $T(Rdm)$. This is just Corollary 5.34.

You should compare the global structure of this section with that of Section 5.2. Once the b&f property is established the development is almost routine.

Exercises

5.7 Write down a set of $\forall_2$-axioms for the theory $T(Rdm)$.

5.8 Show that $V_r \subseteq V_{r+1}$ for each $r < \omega$.

Draw the members of $V_0, V_1, V_2, \ldots$ as finite rooted trees.

5.9 Consider the relation $\sim$ on $V_\omega$. Show that this relation is irreflexive.

5.10 For a given infinite cardinal $\kappa$ let us say a graph $(V, \sim)$ is $\kappa$-random if for each disjoint pair $L, R$ of nodes with $|L|, |R| < \kappa$ we have

$$(\forall x \in L)[x \sim a] \quad (\forall y \in R)[y \sim a]$$

for at least one node $a$.

Consider a set $V$ in the ZF-hierachy. Define $\sim$ on $V$ by

$$x \sim y \iff x \in y \text{ or } y \in x$$

(for $x, y \in V$).

(a) Show that this furnishes $V$ as a graph.

(b) Show that for suitably chosen $V$ this gives a $\kappa$-random graph.
5.11 Let $V$ be a random graph, and let

$$U_1 \cup \cdots \cup U_m$$

be a partition of (the carrier of) $V$ into a finite family of pairwise disjoint (non-empty) sets. We may view each $U_i$ as a subgraph of $V$ (using the restriction to $U_i$ of the mating relation of $V$).

Show that at least one of these subgraphs is random.

5.4 Equivalence relations

In this section we describe a third example of the programme set out in the preamble. This example is, perhaps, not as important as those set out in Section 5.2 and 5.3, but that is not the point. The example is merely an illustration of how general model theoretic methods can work.

5.42 DEFINITION. Let $T(equ)$ and $T(EQU)$ be, respectively, the theory of equivalence relations, and the theory of those equivalence relations where each block is infinite and which have infinitely many blocks.

Thus the models of $T(equ)$ are those structures

$$(A, \approx)$$

where $A$ is a non-empty set and $\approx$ is an equivalence relation, a reflexive, symmetric, transitive relation, on $A$.

This relation $\approx$ partitions $A$ into blocks, the equivalence classes. As usual, for each $a \in A$ we write

$$[a]$$

for the unique block in which $a$ lives. The theory $T(EQU)$ imposes two restrictions on these blocks. Firstly, for each $a \in A$ the block $[a]$ is infinite, there are infinitely many elements $x$ with $a \approx x$. Secondly, there are infinitely many blocks, for each natural number $n$ there are elements

$$a_0, a_1, \ldots, a_n$$

where these elements are pairwise inequivalent.

Thus we have the following.

5.43 THEOREM. The theory $T(equ)$ is $\forall_1$-axiomatizable, and the theory $T(EQU)$ is $\forall_2$-axiomatizable.

As usual we have the two structural properties.

5.44 THEOREM. The theory $T(equ)$ has JEP and AP.

Proof. Dealing with JEP is easy. Given two equivalence structures we may assume they are disjoint, and then we merely take the disjoint union of the carriers with the obvious equivalence relation.

Dealing with AP need just a little more thought.
Consider any wedge of embeddings between equivalence structures.

![Diagram of wedge with A as the common point and B and C as the endpoints connected by f and g, respectively.]

By taking appropriate isomorphic copies we may suppose that f and g are inclusions and

\[ B \cap C = A \]

holds. We now furnish the set

\[ D = (B - A) \cup A \cup (C - A) \]

with an equivalence relation to close the wedge. To do that it is easier to work with the partitioning blocks of the various structures.

Consider any \( a \in A \). This generates blocks

\[ [a]_A \quad [a]_B \quad [a]_C \]

in \( A, B, \) and \( C \). We know that

\[ [a]_B \cap A = [a]_A = [a]_C \cap A \quad [a]_B \cap [a]_C = [a]_A \]

since \( A \) is a substructure of both \( B \) and \( C \), and \( B \cap C = A \). We set

\[ [a]_D = [a]_B \cup [a]_C \]

to obtain a subset of \( D \).

A simple argument show that

\[ [a_1]_D \text{ meets } [a_2]_D \implies [a_1]_D = [a_2]_D \]

for each \( a_1, a_2 \in A \).

This partitions a substantial portion of \( D \), but may be not the whole of \( D \). There may be \( b \in B \) where \([b]_B\) does not meet \( A \). For such a \( b \) we take \([b]_D = [b]_B\). Similarly there may be \( c \in C \) where \([c]_C\) does not meet \( A \). For such a \( c \) we take \([c]_D = [c]_C\).

This partitions the whole of \( D \) and gives us the required equivalence structure. ■

Trivially we have

\[ T(equ) \subseteq T(EQU) \]

but we want a bit more. We need to know that \( T(equ) \) is axiomatized by \( T(EQU) \cap \forall_1 \). This is gives by the following structural property.

5.45 THEOREM. Each equivalence relation is embedable in a big equivalence relation.

Proof. Let \( A \) be an arbitrary equivalence relation. We first enlarge \( A \) be adjoining elements to make each of its blocks infinite. This new equivalence relation may have only finitely many blocks. To deal with this we add infinitely many new blocks each of which is infinite. ■

Now we come to the important part. An analysis of the b&f properties of big equivalence relations.
5.46 Lemma. Let $A, B$ be a pair of big equivalence relations. Let

$$\begin{array}{c}
X
\xrightarrow{f}
Y
\end{array}$$

be a finite partial isomorphism from $A$ to $B$. For each $a \in A$ there is some $b \in B$ and a partial isomorphism

$$\begin{array}{c}
X \cup \{a\}
\xrightarrow{f'}
Y \cup \{b\}
\end{array}$$

which extends $f$.

Proof. We may assume that $a \notin X$, for otherwise we simply take $b = f(a)$.

There are two cases to consider. Suppose $a \approx x$ for some $x \in X$. We require some $b \in (B - Y)$ with $b \approx f(x)$. But the block of $[f(a)]$ of $f(a)$ is infinite and $Y$ is finite, so the difference

$$[f(a)] - Y$$

is non-empty, and any element $b$ of this set will do.

Suppose $a \approx x$ for no $x \in X$. We require some $b \in (B - Y)$ with $b \not\approx f(y)$ for each $y \in Y$. Since $B$ has infinitely many blocks, there is such an element $b$. ■

In the standard way we may use this construction in both directions to obtain the following.

5.47 Theorem. For each pair $A, B$ of big equivalence relations, the family of all finite partial isomorphisms from $A$ to $B$ is a b&f system.

With this we can very quickly verify the required properties (0, 1, 2, 3) of $T(EQU)$.

5.48 Theorem. For each pair $A, B$ of big equivalence relations we have

$$A \equiv_p B$$

and, in particular, the theory $T(EQU)$ is $\aleph_0$-categorical and complete.

Next we need to verify model completeness. For that we again invoke the b&f property.

5.49 Lemma. Let $A, B$ be a pair of big equivalence relations. Then we have

$$(A, a) \equiv_0 (B, b) \implies (A, a) \equiv_p (B, b)$$

for each pair $(a, b)$ finite sequences from $A$ and $B$.

Proof. Suppose

$$(A, a) \equiv_0 (B, b)$$

and let $X \subseteq A$ and $Y \subseteq B$ be these finite sequences viewed as sub-equivalence structures of $A$ and $B$. Thus we have a partial isomorphism

$$\begin{array}{c}
X
\xrightarrow{f}
Y
\end{array}$$

from $A$ to $B$. By Theorem 5.47 this is the base of a b&f system between the two structures $(A, a)$ and $(B, b)$. Thus

$$(A, a) \equiv_p (B, b)$$

as required. ■

Looking at this in a slightly different way gives the following.
5.50 **THEOREM.** The theory $T(EQU)$ is model complete.

**Proof.** Consider any pair of big equivalence relations

$$A \subseteq B$$

one included in the other. Consider any finite list $a$ of elements of $A$. We have

$$(A, a) \subseteq (B, b)$$

so that

$$(A, a) \equiv_0 (B, b)$$

and hence

$$(A, a) \equiv_p (B, b)$$

by Lemma 5.49 with

$$(A, a) \equiv (B, b)$$

as a special case. By letting $a$ vary through all possible finite sequences we have

$$A \prec B$$

as required. ■

Earlier we verified that $T(equ)$ has AP, so we may strengthen Theorem 5.50.

5.51 **THEOREM.** The theory $T(Line)$ has EQ.

**Proof.** By Theorems 4.10 and 5.50 it suffices to show that the theory (axiomatized by)

$$T(Line) \cap \forall_1$$

has AP. By Theorem 5.45 this theory is just $T(equ)$, and by Theorem 5.44 this theory has AP. ■

With this we have completed another pass through the general programme.

5.52 **THEOREM.** The theory $T(EQU)$ is a ‘canonical’ companion of $T(equ)$.

**Proof.** We have $T(equ) \subseteq T(EQU)$ where $T(EQU)$ has the required properties (0, 1, 2, 3). Thus it suffices to show that each model of $T(equ)$ is embeddable in some model of $T(EQU)$. This is just Theorem 5.45. ■

Once again you should observe the global structure of this section. The b&f property gives us almost everything.

**Exercises**

5.12 Complete the proof of Theorem 5.44.
5.5 Miniature arithmetic

In this section we look at another example of the programme set out in the preamble, but with a slightly different emphasis. We don’t first produce a minor theory \( S \) and then a major theory \( T \) (which turns out to be a ‘canonical’ companion of \( S \)). We concentrate on just one theory \( T \) and show this has the properties \((0, 1, 2, 3)\). We do look at the theory \( S \) axiomatized by \( T \cap \forall \), but it turns out that this has no intrinsic interest.

In Section 2.3 we produced a rather simple axiomatization of

\[
Th(\mathbb{N}, S, 0)
\]

and showed that this theory has \( EQ \). We did that by actually describing an algorithm which eliminates quantifiers from formulas. Unfortunately, that analysis didn’t tell us very much about the theory and its class of models. In this section we look at an enrichment

\[
Th(\mathbb{N}, \leq, S, 0)
\]

of that theory obtained by adjoining a symbol for the linear comparison on \( \mathbb{N} \). We show how model theoretic methods give us much more information.

5.53 DEFINITION. In the language with one 2-placed relation symbol \( \leq \), one 1-placed operation symbol \( S \), and one constant 0, consider the following set of axioms.

\begin{align*}
(L1) \quad & (\forall v)[v \leq v] \\
(L2) \quad & (\forall u, v, y)[u \leq v \leq w \rightarrow u \leq w] \\
(L3) \quad & (\forall u, v)[u \leq v \leq u \rightarrow u \equiv v] \\
(L4) \quad & (\forall u, v)[u \leq v \lor v \leq u]
\end{align*}

\begin{align*}
(S0) \quad & (\forall v)[Sv \neq 0] \\
(S1) \quad & (\forall u, v)[Su \equiv Sv \rightarrow u \equiv v] \\
(S2) \quad & (\forall v)[v \neq 0 \rightarrow (\exists u)[Su = v]] \\
(S3_k) \quad & (\forall v)[S^{k+1} v \neq v] \quad \text{(for all } k \in \mathbb{N})
\end{align*}

\begin{align*}
(I) \quad & (\forall v)[v \leq Sv] \\
(M) \quad & (\forall u, v)[u \leq v \rightarrow Su \leq Sv] \\
(Z) \quad & (\forall v)[0 \leq v] \\
(D) \quad & (\forall u, v)[u \leq Sv \rightarrow u \leq v \lor u = Sv]
\end{align*}

Let

\[
T(\leq, S, 0)
\]

be the theory axiomatized by this set of sentences.
Let’s have a look at these axioms in a bit more detail.

The collection \((L1 – L4)\) axiomatize linearly ordered sets. That theory and its models are discussed in Section 5.2. In this section we make use of the arithmetic of linear order types.

The axioms \((S1 – S3)\) are lifted from Definition 2.9. As we saw in Section 2.3 these axiomatize the theory

\[ Th(\mathbb{N}, S, 0) \]

and this has \(EQ\).

Here we are interested in the interaction of these two notions. In particular, the theory

\[ Th(\mathbb{N}, \leq, S, 0) \]

is a primary interest. To investigate that theory we add the four mixed axioms \((I, M, D, Z)\). Of these axioms \((I, M)\) say that the operation \(S\) is inflationary and monotone, and \((Z)\) says that 0 is the least member of the linear ordering. As we will see later, axiom \((D)\) ensures the ordering is discrete.

Notice that almost all these axioms are \(\forall_1\)-sentences. The only one that isn’t is \((S2)\), and this is an \(\forall_2\)-sentences. Thus, by construction, the theory \(T(\leq, S, 0)\) is \(\forall_2\)-axiomatizable. I don’t know if these are a minimal collection of axioms for the theory.

Trivially, as they say, we have

\[ T(\leq, S, 0) \subseteq Th(\mathbb{N}, \leq, S, 0) \]

since \(\mathbb{N}\) (with the indicated furnishings) is a model of \(T(\leq, S, 0)\). In due course we show that these two theories are equal.

We know that \(\mathbb{N}\) can not be the only model of \(T(\leq, S, 0)\) (for the theory must have uncountable models) but perhaps it is the only countable model. I’m afraid not, but as compensation we can give a complete description of all models.

Consider any element \(a\) of a model of \(T(\leq, S, 0)\). By \((I)\) we have \(a \leq Sa\) and \((S3_0)\) ensures \(a < Sa\). Is there anything between these two elements?

5.54 LEMMA. For elements \(a, b\) of an arbitrary model of \(T(\leq, S, 0)\) we have

\[ a \leq b \leq Sa \implies a = b \text{ or } b = Sa \]

and hence \(a < Sa\) is a gap.

Proof. Assuming

\[ a \leq b \leq Sa \]

a use of \((D)\) gives

\[ a \leq b \leq a \text{ or } b = Sa \]

and then \((L3)\) gives the required result. ■

With this observation we can quickly locate the structure of an arbitrary model.

5.55 THEOREM. Each model of \(T(\leq, S, 0)\) consists of \((a \text{ copy of }) \mathbb{N}\) followed by several copies of \(\mathbb{Z}\) arranged in a linear sequence (where both \(\mathbb{N}\) and \(\mathbb{Z}\) carry the obvious furnishings).
**Proof.** Let $A$ be any model of $T(\leq, S, 0)$ and consider any $a \in A$. Using (I) we generate a sequence

$$a \leq Sa \leq S^2a \leq \cdots \leq S^r a \leq \cdots$$

through $A$. We check that this sequence doesn’t cycle.

Suppose $S^r a = S^s a$

for some $r \leq s$ (from $\mathbb{N}$). By repeated use of (S1) we obtain

$$S^t a = a$$

where $t = s - r$. But now (S3) gives $t = 0$, and hence $r = s$.

This shows that each element $a$ generates a copy of $\mathbb{N}$ in $A$. By Lemma 5.54 there is nothing else within the range of this copy. In particular, with $a = 0$, axiom (Z) ensures that $A$ starts with (a copy of) $\mathbb{N}$ and everything else in $A$ is beyond this initial part.

Consider any non-standard element $a$ of $A$, that is any element $a$ that is not in this initial stretch. By repeated use of the crucial axiom (S2) we may generate a sequence $a(\cdot)$ of elements by

$$a(0) = a \quad Sa(r + 1) = a(r)$$

for each $r \in \mathbb{N}$. We check that this sequence doesn’t cycle.

Suppose $a(s) = a(r)$

for some $r \leq s$. With $t = s - r$ we have

$$S^t a(s) = a(r) = a(s)$$

and hence (S3) gives $t = 0$ and $r = s$.

This shows that each non-standard element $a$ generates a copy on $\mathbb{Z}$

$$\cdots < S^{-2}a < S^{-1}a < a < Sa < S^2a < \cdots$$

and Lemma 5.54 ensures that there is nothing else within this two-way stretch.

This shows that $A$ consists of $\mathbb{N}$ together with several non-overlapping copies of $\mathbb{Z}$. Since the whole of $A$ is linearly ordered, these non-standard blocks are arranged in a linear sequence. 

Using linear order types we can rephrase Theorem 5.55 in a succinct way.

5.56 **COROLLARY.** Each model $A$ of $T(\leq, S, 0)$ can be uniquely decomposed as

$$\mathbb{N} + \mathbb{Z} \cdot \alpha$$

where $\alpha$ is a linear order type.

This shows that in some ways the analysis of $T(\leq, S, 0)$ is nothing more than a analysis of linear order types.

In this description $\alpha$ can be any order type whatsoever (including the empty type). As a particular case we make take the rationals for $\alpha$ to obtain a rather special model

$$\mathcal{Q} = \mathbb{N} + \mathbb{Z} \cdot \mathbb{Q}$$
of $T(\leq, S, 0)$. This has an an important role to play. In particular, later we will see that

$$\mathbb{N} \prec \Omega$$

which is not at all obvious.

By the manipulation of order types we can obtain a structural property of $T(\leq, S, 0)$.

5.57 THEOREM. The theory (axiomatized by)

$$T(\leq, S, 0) \cap \forall_1$$

has $AP$.

Proof. As a preliminary consider any submodel $M$ of $T(\leq, S, 0)$. Thus $M$ is a substructure of at least one model $A$ of $T(\leq, S, 0)$. What bits of $A$ can be missing in $M$? The structure $M$ must contain 0 and be closed under $S$. Thus either some of the $\mathbb{Z}$-blocks of $A$ have the left hand end missing, or the whole of some $\mathbb{Z}$-blocks have been omitted. This shows there is a canonical, minimal, model $A$ extending $M$. We take that model in which each $\mathbb{Z}$-block is represented in $M$. We can reconstruct from $M$ by filling in the missing left hand ends of the short $\mathbb{Z}$-blocks.

Now consider a wedge

of submodels. This gives us a wedge of models

where each is the canonical extension of its submodel. We need to describe the two embeddings $l$ and $r$. By symmetry we can deal with $l$.

Consider any element $a \in A$. If $a$ is standard then $a$ is in $M$ and so $l(a)$ is determined by the image of $a$ in $L$. If $a$ is non-standard then it occurs in a unique non-standard block of $A$, and can be viewed as $m$ in that block for some unique $m \in \mathbb{Z}$. Although $m$ may not appear in $M$, this block is certainly represented in $M$, and past to a partial block in $L$, which generates a whole block in $B$. This block of $B$ has its own $m$, an we take that as $l(a)$.

This shows that the larger diagram does commute, and so it suffices to close the upper wedge of models.

We construct a model $D$ and a pair of embeddings such that the square commutes.
By replacing \( B, C \) by isomorphic copies we may suppose that \( l, r \) are insertions and that
\[
B \cap C = A
\]
holds. Thus we may set
\[
D = A \cup (B - A) \cup (C - A)
\]
and it doesn’t matter how we order the blocks that are added to \( A \). \[\Box\]

You may wonder what the theory
\[
T(\leq, S, 0) \cap \forall_1
\]
is. In fact, by perusing the proofs so far we see that it is the theory axiomatized by the sentences \( (L1 - L4, S0, S1, S3_k, I, M, Z, D) \). Thus we simply omit the only \( \forall_2 \)-axiom \( S2 \) from \( T(\leq, S, 0) \).

Next we begin to analyse the embeddings between models of \( T(\leq, S, 0) \).

5.58 DEFINITION. For non-standard elements \( l, r \) of a model \( A \) of \( T(\leq, S, 0) \) we write
\[
l \ll r
\]
and say \( l \) and \( r \) are separated (in that order) if \( l \) and \( r \) lie in different blocks and the block of \( l \) is to the left of that of \( r \), that is if
\[
l + n \leq r
\]
for each \( n \in \mathbb{N} \). \[\Box\]

In this definition we have written
\[
l + n \text{ for } S^n l
\]
which is a more suggestive notation.

5.59 LEMMA. Let \( A \) be any model of \( T(\leq, S, 0) \) and let \( l \ll r \) be a separated pair of non-standard elements. Then there is an elementary extension \( A \prec B \) with
\[
x \ll l \ll y \ll r \ll z
\]
for some elements \( x, y, z \) of \( B \).

Proof. Let \( a \) be an enumeration of \( A \) and enrich the language by adding these as constants. Amongst these will be (names for) \( l \) and \( r \). We also add three more constants \( x, y, z \). We are looking for a model of
\[
Th(A, a) \cup \Sigma
\]
for a certain set \( \Sigma \) of sentences in the enriched language.

Let \( \Sigma \) be the set of all sentences
\[
x + n \leq l \quad l + n \leq y \quad y + n \leq r \quad r + n \leq z
\]
for all $n \in \mathbb{N}$. Let $\Gamma$ be any finite subset of $\Sigma$. It suffices to show that

$$Th(A, a) \cup \Gamma$$

has a model. We show that when $x, y, z$ are suitably interpreted the structure $A$ provides such a model.

Within $\Gamma$ there is a largest $n$ that occurs. Consider the block in which $l$ lives. This extends both ways. We take $x, y$ in that block with $x$ sufficiently small and $y$ sufficiently large. This ensures

$$x + n \leq l \quad l + n \leq y$$

and then

$$y + n \leq r$$

since, in fact, $y \ll r$. Finally, we satisfy

$$r + n \leq z$$

by taking $z$ to be a sufficiently large many of the block of $r$. 

We need a more powerful version of this result.

5.60 Lemma. Let $A$ be any model of $T(\leq, S, 0)$. Then there is an elementary extension $A \prec B$ with $|A| = |B|$, and such that for each pair $l \ll r$ of elements of $A$ we have

$$x \ll l \ll y \ll r \ll z$$

for some elements $x, y, z$ of $B$ (which may depend on $l, r$).

Proof. We refine the proof of Lemma 5.59.

First of all we enrich the language by adding a list $a$ of constants to name the elements of $A$. Next for each pair $l \ll r$ from $A$ we add three more constants

$$x(l, r) \quad y(l, r) \quad z(l, r)$$

with a different triple for each pair. Observe that the full enriched language has cardinality $|A|$.

For each pair $l \ll r$ let $\Sigma(l, r)$ be the set of all sentences

$$x(l, r) + n \leq l \quad l + n \leq y(l, r) \quad y(l, r) + n \leq r \quad r + n \leq z(l, r)$$

for all $n \in \mathbb{N}$. Let $\Sigma$ be the union of all these sets $\Sigma(l, r)$. By modifying the argument of the proof of Lemma 5.59 we see that

$$Th(A, a) \cup \Sigma$$

is finitely satisfiable in $A$, and hence the required model exists. 

Recall that each countable linear order can be embedded in $\mathbb{Q}$. Recall also that we have a special model

$$\mathbb{Q} = \mathbb{N} + \mathbb{Z} \cdot \mathbb{Q}$$

of $T(\leq, S, 0)$. We combine these two observations.
5.61 **Theorem.** Each countable model of $T(\leq, S, 0)$ is elementarily embeddable in the special model $\mathcal{Q}$.

**Proof.** Consider any countable model $A$. By Lemma 5.60 there is a countable elementary extension $A \prec A'$ with the property that for each pair $l \ll r$ of elements of $A$ we have

$$x \ll l \ll y \ll r \ll z$$

for some elements $x, y, z$ of $A'$.

We iterate this construction to produce an ascending chain

$$A \prec A' \prec A'' \prec \cdots \prec A^{(i)} \prec \cdots$$

of models. Let $B$ be the union of the chain. This $B$ is a countable elementary extension of $A$. It suffices to determine the order type of its non-standard blocks.

Consider any two non-standard element $l \ll r$ of $B$. These are members of some model $A^{(i)}$. By construction we have

$$x \ll l \ll y \ll r \ll z$$

for some $x, y, z$ in $A^{(i+1)}$. This shows that the non-standard blocks of $B$ are densely ordered without end points, and hence have order type $\mathbb{Q}$. ■

A simple use of the downwards L-S technique now gives the following.

5.62 **Theorem.** The theory $T(\leq, S, 0)$ is complete.

In particular, the $\forall_2$-axiomatizable theory $T(\leq, S, 0)$ is nothing more than the theory $Th(N, \leq, S, 0)$ as we claimed earlier.

The crucial trick of the section is to produce a b&f system on the special model $\mathcal{Q}$. In fact, we use the finitely generated partial automorphisms, but we need to handle these with a bit of care.

5.63 **Definition.** A model $A$ of $T(\leq, S, 0)$ has finite width if it has the form

$$A = \mathbb{N} + \mathbb{Z} + \cdots + \mathbb{Z}$$

where there are just finitely many non-standard blocks. ■

We know that each model of finite width can be embedded in $\mathcal{Q}$. In fact, it can be embedded in $\mathcal{Q}$ in many different ways (except if it is the standard model $\mathbb{N}$). We need a way of indexing these embeddings.

Let

$$I = (0, 1) \cap \mathbb{Q}$$

the set of rational numbers $q$ with $0 < q < 1$. Of course, as a linearly ordered set $I$ and $\mathbb{Q}$ are isomorphic, so we may use $I$ to index the non-standard blocks of $\mathcal{Q}$. We let

$$\mathcal{Q} = \mathbb{N} + \{\mathbb{Z}_i | i \in I\}$$
where the non-standard blocks are ordered according to their indexes.

Consider any finite strictly increasing list
\[ i ::= i(1) < i(2) < \cdots < i(l) \]
of indexes from \( I \). Each such list \( i \) determines a unique substructure of \( \Omega \)
\[ A_i ::= \mathbb{N} + \mathbb{Z}_{i(1)} + \cdots + \mathbb{Z}_{i(l)} \]
formed by using those blocks \( \mathbb{Z} \) indexed in \( i \). The empty list \( \bot \) is allowed, in which case
\[ A_\bot = \mathbb{N} \]
the minimal model.

Consider two such list \( i, j \) with the corresponding substructures
\[ A_i, A_j \]
uniquely determined by the two lists. These two structures are isomorphic precisely when \( i \) and \( j \) have the same length.

Consider two such structures
\[ A_i ::= \mathbb{N} + \mathbb{Z}_{i(1)} + \cdots + \mathbb{Z}_{i(l)} \]
\[ A_j ::= \mathbb{N} + \mathbb{Z}_{j(1)} + \cdots + \mathbb{Z}_{j(l)} \]
using lists of equal length. What does an isomorphism between the two look like? It has no choice on the standard part. It must then match
\[ \mathbb{Z}_{i(1)} \text{ with } \mathbb{Z}_{j(1)} \]
\[ \cdots \]
\[ \mathbb{Z}_{i(l)} \text{ with } \mathbb{Z}_{j(l)} \]
in some way. When doing this matching it has a choice, because there is no distinguished element in each block. There are many such matchings and for what we do here we need not be more specific.

This gives us a family of partial isomorphisms from \( \Omega \) to \( \Omega \). Each has the form
\[ A_i \xrightarrow{(i,j)} A_j \]
where \( i \) and \( j \) are list of the same length. This notation doesn’t uniquely determined the isomorphism but, as remarked, we need to be more specific. For convenience let us say such a partial isomorphism has finite width. Observe that, in fact, these are just the finitely generated partial isomorphisms from \( \Omega \) to \( \Omega \).

5.64 LEMMA. The family of all partial isomorphisms of finite width is a b\&f system on \( \Omega \).

Proof. By symmetry it suffices to verify just one of the two extension properties. Let us go forth.

Consider one of the selected partial isomorphisms
\[ A_i \xrightarrow{(i,j)} A_j \]
where both \( i \) and \( j \) have length \( l \).
5.5. Miniature arithmetic

Consider any element $a$ of $\Omega$ viewed as a superstructure of $A_i$. If $a$ lives in $A_i$, then we have nothing to do.

Otherwise $a$ lives in some block $Z_p$ for some unique $p \in I$. Compare $p$ with the indexes

$$i(1) < \cdots < i(l)$$

in $i$. The new index $p$ lives inside exactly one of the intervals of $I$ determined by the list (either on the extreme left, or the extreme right, or between two consecutive indexes in $i$). By inserting $p$ in the appropriate position we obtain a list $i^+$ with one extra member (namely $p$).

We note the interval in which $p$ lives and go to the corresponding interval in $j$. We select any rational $q$ in that interval, insert it into $j$ to obtain a list $j^+$ with one extra member (namely $q$).

By matching the two new blocks

$$A_i \xrightarrow{(i^+, j^+)} A_j$$

we obtain an appropriate extension.  

Since $\Omega$ is countable this b&f system ensures that $\Omega$ is homogeneous.

5.65 COROLLARY. On $\Omega$ each partial isomorphism of finite width $\Omega$ lifts to a full automorphism.

By Theorem 5.61 for each countable model there is at least one elementary embedding into $\Omega$. But what about other embeddings? Are these also elementary? The homogeneity of $\Omega$ given by Corollary 5.65 ensures a positive answer.

First we prove a special case of this result.

Observe that the following sequence of arguments has a rather general nature, and hasn’t much to do with this particular theory.

5.66 LEMMA. Consider any embedding $A \subseteq \Omega$ where $A$ is a model of finite width. Then $A \prec \Omega$.

Proof. We are given $A \subseteq \Omega$ and by Theorem 5.61 we know that $A$ has an elementary embedding into $\Omega$. However, that particular embedding may not be the one we are given here. By replacing $A$ by a suitable isomorphic copy we obtain a diagram

$$\Omega \xrightarrow{f} \Omega$$

where $f$ is an isomorphism and the right hand embedding is elementary, as indicated. By Lemma 5.64 this isomorphism $f$ is a member of a b&f system on $\Omega$, and hence by
Corollary 5.65 lifts to a full automorphism of $\Omega$. Thus we have a commuting square

\[
\begin{array}{ccc}
\Omega & \xrightarrow{g} & \Omega \\
\downarrow{l} & & \downarrow{r} \\
A & \xrightarrow{f} & A'
\end{array}
\]

where $g$ is an isomorphism. To help with the following argument we have labelled the two embeddings.

Consider any formula $\phi(v)$ and point $a$ of $A$ matching the batch $v$. Going round the second square we have

\[
A \models \phi(a) \implies A' \models \phi(f(a)) \\
\implies \Omega \models \phi((r \circ f)(a)) \\
\implies \Omega \models \phi((g \circ l)(a)) \implies \Omega \models \phi(l(a))
\]

for the required result. At the first and last step we use the passage across an isomorphism, the second step uses the given elementary embedding, and the third step uses the commuting square.

We now improve this in a series of steps.

5.67 LEMMA. We have

\[
B \subseteq \Omega \implies B \prec_1 \Omega
\]

for each countable model $B$ of $T(\leq, S, 0)$.

Proof. Suppose $B \subseteq \Omega$ and that

\[
B \models \phi(a)
\]

for some $\forall_1$-formula $\phi(v)$ and point $a$ from $B$ that matches the batch $v$. There are only finitely many elements in $a$, and these live in some $A \subseteq B$ of finite width. (We simply select those non-standard blocks of $B$ that are occupied by at least one member of $a$.) This gives us

\[
A \subseteq B \subseteq \Omega
\]

where the elementary embedding holds by Lemma 5.66. Since $\phi$ is $\forall_1$ we have

\[
A \models \phi(a)
\]

and then the elementary embedding gives

\[
\Omega \models \phi(a)
\]

for the required result.

With this we can deal with pairs of countable models.
5.68 **Lemma.** We have $$B \subseteq C \implies B \prec_1 C$$ for each pair $B, C$ of countable models of $T(\leq, S, 0)$.

**Proof.** For a pair $B \subseteq C$ a use of Theorem 5.61 (in a weakened form) and Lemma 5.67 gives us

$$B \subseteq C \subseteq \Omega \xrightarrow{\prec_1}$$

and then $B \prec_1 C$ follows by a standard argument. 

With this a standard interlacing argument gives the following.

5.69 **Corollary.** We have $$B \subseteq C \implies B \prec C$$ for each pair $B, C$ of countable models of $T(\leq, S, 0)$.

This deals with the countable models. Now we have to deal with all models. There are various ways of doing that. Here is one way. We generalize Lemma 5.68.

5.70 **Lemma.** We have $$B \subseteq C \implies B \prec_1 C$$ for each pair $B, C$ of models of $T(\leq, S, 0)$.

**Proof.** Suppose $B \subseteq C$ and suppose

$$C \models (\exists v)\delta(b, v)$$

where $\delta(u, v)$ is a quantifier-free formula and $b$ is a point from $B$ that matches the batch $u$. We have

$$C \models \delta(b, c)$$

for some point $c$ of $C$ which matches the batch $v$.

Let $B' \subseteq B$ be a countable model which contains $b$. Let $C' \subseteq C$ be a countable model that includes $B$ and contains $c$. (For this theory these are easy to produce. We simply select the appropriate finite number of non-standard blocks of $B$ and $C$.) This gives us a commuting square

$$
\begin{array}{ccc}
B & \prec & C \\
\uparrow & & \uparrow \\
B' & \prec & C'
\end{array}
$$

where the lower elementary embedding holds by Corollary 5.69. (As we will see, Lemma 5.68 suffices for the following argument.)

Remembering that $\delta$ is quantifier-free and $c$ is in $C'$ we have

$$C' \models \delta(b, c)$$
so that

\[ C' \models (\exists v)\delta(b, v) \]

to give

\[ B' \models (\exists v)\delta(b, v) \]

and hence

\[ B \models (\exists v)\delta(b, v) \]

for the required result.

With this a standard argument (as for the proof of Corollary 5.69) gives the following.

**5.71 THEOREM.** The theory \( T(\leq, S, 0) \) is model complete.

Finally, we remember the characterization of Theorem 4.10 together with Theorem 5.57 to obtain the following.

**5.72 THEOREM.** The theory \( T(\leq, S, 0) \) has EQ.

It is possible to extend the algorithm of Section 2.3 to give an explicit descriptions of the quantifier elimination procedure for \( T(\leq, S, 0) \). However, that algorithm doesn’t tell us very much. The model theoretic analysis provides much more information about the models of \( T(\leq, S, 0) \).

**Exercises**

5.13 Show that the theory \( T(\leq, S, 0) \cap \forall_1 \) has JEP.

[Need some more]

5.6 Perhaps algebraically closed abelian groups if I can remember the details

5.7 Survey – of other examples
6

Companion theories and existentially closed structures

Model complete theories are nice, but not all theories are model complete. We know that if a theory is model complete then each extension of that theory (in the same language) is model complete. For any language $L$ the smallest theory – the pure logic – the theory of all $L$-structures, is not model complete. If by some devilish trick this pure logic turned out the be model complete the the manipulation of quantifiers in that language would be almost trivial. And we can’t have that can we!

A theory $T$ need not be model complete, but maybe we can move, in a canonical fashion, from $T$ to some model complete theory $T^*$ (in the same language), and one which has a close relationship with $T$. There are such theories $T$, and then the associated theory $T^*$ is central in the analysis of $T$. A few of the simpler examples are developed in Chapter 5. In this chapter most make precise the informal notion of the ‘canonical’ companion used in Chapter 5.

Before we start the development in earnest it is worth mentioning, but not analysing, a more complicated example.

Let $T$ be theory of fields (perhaps with specified characteristic), and let $T^*$ the theory of algebraically closed fields (of that characteristic). Every field can be embedded in an algebraically closed field, in fact it has an algebraic closure. By pursuing these ideas we find that $T^*$ is model complete, and in fact has $EQ$. An analysis of the relationship between $T$ and $T^*$, the relationship between general fields and algebraically closed fields, was an important initial step in the development of model theory.

Returning to the general case, suppose for the theory $T$ under investigation there is no suitable model complete theory $T^*$ associated with $T$. What do we do for such a case?

There is two kinds of possibilities.

We can relax the required syntactic conditions of the associated theory.

Or we can move to a non-elementary class which has some, but not all, of the properties of the models of a model complete theory.

In this chapter we begin to develop these ideas.

6.1 Model companions

Given a theory $T$ where might we look for a model complete theory $T^*$ which is closely attached to $T$? We need to make precise what ‘closely attached’ should mean.

6.1 DEFINITION. Two theories $T_1, T_2$ (in the same language) are companions if each model of the one can be embedded in a model of the other.

Thus, two theories $T_1, T_2$ are companions exactly when $S(T_1) = S(T_2)$ or, equivalently, when $T_1 \cap \forall_1 = T_2 \cap \forall_1$. For each language this puts an equivalence relation on the family of all theories in that language. We look for special members of each companion block.
6.2 LEMMA. Suppose $T$ and $T^*$ are companion theories with $T^*$ model complete. Then $T \cap \forall_2 \subseteq T^*$.

Proof. Consider any $A \models T^*$. We require $A \models T \cap \forall_2$.

Two uses of the companion property gives

$$A \subseteq B \subseteq A' \quad A, A' \models T^* \quad B \models T$$

and then $A \prec A'$ since $T^*$ is model complete. But now $A \prec_1 B$, and hence $A \models T \cap \forall_2$, as required. 

We now remember that each model complete theory is $\forall_2$-axiomatizable, and so obtain the following.

6.3 COROLLARY. Suppose $T_1$ and $T_2$ are model complete companion theories. Then $T_1 = T_2$.

This shows that each companion block contains at most one model complete theory. We select this nice member when it exists, and a good approximation when it doesn’t.

6.4 DEFINITION. A model companion of a theory $T$ is a companion $T^*$ of $T$ which is model complete.

Corollary 6.3 shows that each theory has at most one model companion. Some theories do have a model companion and some don’t. Let’s look at some examples.

6.5 EXAMPLE. Let $T$ be any one of the three theories considered in Sections 5.2, 5.3, 5.4. In each case we produced a theory $T^*$ where $T^*$ has $E Q$ with $T \subseteq T^*$ and where each model of $T$ is embeddable in a model of $T^*$. Thus, in each case, $T^*$ is the model companion of $T$.

This gives us three examples of a theory $T$ with a model companion $T^*$. In each case the theory $T^*$ has rather better properties than just being model complete. Of course, not all model companions are like this.

Next we go to the other extreme and look at a theory without a model companion. This will take a little longer to develop, but illustrates many aspects of this topic.

6.6 EXAMPLE. Let $T$ be the theory of commutative rings with 1. Such a ring is viewed as a structure $A = (A, +, \times, 0, 1)$ where $0 \neq 1$. We follow the common practice and identify each ring $A$ with its carrier $A$. Also we may drop the modifier ‘commutative’ since every ring we meet is commutative.

We show that $T$ does not have a model companion. We argue by contradiction. Thus we assume that $T$ has a model companion $T^*$ and eventually reach a contradiction.

The theory $T$ is $\forall_2$-axiomatizable. Most of the axioms are $\forall_1$-sentences, but the existence of additive inverses requires a $\forall_2$-axiom. In particular $T \subseteq T^*$ holds. To obtain the contradiction we need a trick.
Recall that an element $a$ of a ring $A$ is nilpotent if $a^n = 0$ for some $n \in \mathbb{N}$. This can be expressed by an infinite disjunction
\[
\bigvee \{v^n = 0 \mid n \in \mathbb{N}\}
\]
and we show that, modulo any model companion $T^*$, this can be reduced to a single formula. Let $\text{nnil}(v)$ be the $\exists_1$-formula
\[
(\exists w)((vw \not= 0) \land ((vw)^2 = vw))
\]
where $v$ is some selected variable. This $\exists_1$-formula has many quantifier-free consequences.

6.7 LEMMA. For each $n \in \mathbb{N}$
\[T \vdash \text{nnil}(v) \rightarrow (v^n \not= 0)\]
holds.

This is proved by a simple calculation in an arbitrary ring. A similar calculation is used in the following crucial observation.

6.8 LEMMA. For each ring $A$ and element $a$ of $A$, the following are equivalent.

(i) The element $a$ is not nilpotent.

(ii) The is some ring $A \subseteq B$ with $B \models \text{nnil}(a)$.

Proof. $(i) \Rightarrow (ii)$. Suppose $a$ is not nilpotent and consider the factor ring
\[B = A[X]/\langle aX - a^2X^2 \rangle\]
of the polynomial extension $A[X]$ of $A$. A typical member of the ideal $\langle aX - a^2X^2 \rangle$ has the form
\[p(X) = (b_0 + b_1X + b_2X^2 + \cdots)(aX - a^2X^2) = c_0 + c_1X + c_2X^2 + c_3 + X^3 + \cdots\]
where $b_0, b_1, b_2, \ldots$ is a list of elements of $A$ with $b_n = 0$ for all sufficiently large $n$. On multiplying out we see that
\[
c_0 = 0
\]
\[
c_1 = ab_0
\]
\[
c_2 = ab_1 - a^2b_0
\]
\[
\vdots
\]
\[
c_{n+2} = ab_{n+1} - a^2b_n
\]
\[
\vdots
\]
are the coefficients of $p(X)$. In particular, the only member of $A$ that belongs to the ideal is 0, and hence the canonical morphism
\[A \longrightarrow B\]
is an embedding.
It now suffices to show that \( aX \notin \langle aX - a^2X^2 \rangle \).

By way of contradiction, suppose \( aX \in \langle aX - a^2X^2 \rangle \). Then, using the case \( p(X) = aX \) we have

\[
\begin{align*}
    a &= ab_0 \\
    0 &= ab_1 - a^2b_0 \\
    &\vdots \\
    0 &= ab_{n+1} - a^2b_n
\end{align*}
\]

for some eventually constant list \( b_0, b_1, b_2, \ldots \) of elements of \( A \). A simple induction gives

\[
a^{n+1} = ab_n
\]

(for \( n \in \mathbb{N} \)), and and hence \( a^{n+1} = 0 \) for all large \( n \). This is the contradiction, since \( a \) is not nilpotent.

(ii) \( \Rightarrow \) (i). This follows by a simple calculation. \( \blacksquare \)

This embedding result quickly leads to the required result.

6.9 THEOREM. The theory \( T \) of commutative rings with 1 does not have a model companion.

Proof. By way of contradiction, suppose \( T^* \) is the model companion of \( T \). We show that

\[
T^* \cup \{ (v^n \neq 0) \mid n \in \mathbb{N} \} \vdash \text{nnil}(v)
\]

holds.

To this end, consider any model \( (A,a) \) of the hypothesis set. Thus \( A \models T^* \), and \( a \) is an element of \( A \) which is not nilpotent. By Lemma 6.8, there is some \( A \subseteq B \models T \) with \( b \models \text{nnil}(a) \). There is some \( B \subseteq C \models T^* \), and then \( C \models \text{nnil}(a) \) since \( \text{nnil}(v) \) is a \( \exists_1 \)-formula. But \( T^* \) is model complete, so that \( A \prec C \), and hence \( A \models \text{nnil}(a) \), as required.

We now apply compactness to get

\[
T^* \vdash (v^n \neq 0) \rightarrow \text{nnil}(v)
\]

and hence

\[
T^* \vdash (\forall v)[(v^n \neq 0) \rightarrow \text{nnil}(v)]
\]

for some \( n \in \mathbb{N} \). But now, using Lemma 6.7, we have

\[
T^* \vdash (\forall v)[(v^n \neq 0) \rightarrow (v^{n+1} \neq 0)]
\]

and hence

\[
T \vdash (\forall v)[(v^n \neq 0) \rightarrow (v^{n+1} \neq 0)]
\]

since we are dealing with a \( \forall_1 \)-sentence. This is the contradiction since, for each \( n \), it is easy to construct a ring with an element \( a \) where \( a^{n+1} = 0 \) but \( a^n \neq 0 \). \( \blacksquare \)

This completes the development of Example 6.6.

These examples and ideas pose several questions.
6.2. Companion operators

When does a theory have a model companion? What special properties does this companion have?

When a theory does not have a companion, is there any kind of replacement which is nearly as good?

In this chapter we investigate these and similar questions.

The notion of a model companion of a theory was refined (by a rather interesting gentleman, Eli Bers) from an earlier notion of a model completion. It is worth looking at this ancestor.

6.10 DEFINITION. A model completion of a theory $T$ is a companion $T^*$ such that $T \subseteq T^*$ and such that $T^*[\mathfrak{A}]$ is complete for each model $\mathfrak{A} \models T$.

Exercise 6.1 indicates how this relates to the notion of a model companion and why model companions are more useful than completions.

Exercises

6.1 (a) Suppose $T^*$ is a model completion of a theory $T$. Show that $T^*$ is the model companion of $T$, and $T$ has AP.

(b) Suppose $T$ has AP and a model companion $T \subseteq T^*$. Show that $T^*$ is a model completion of $T$.

(c) Find examples of theories which do have a model completion.

6.2 Complete the proofs of Lemmas 6.7 and 6.8.

6.2 Companion operators

What can we do when a theory $T$ does not have a model companion? We need some kind of substitute. We can view the attempted selection of a model companion in two ways, both of which lead to different possible substitutes.

- When the theory $T$ has a model companion, we attach to $T$ a particularly nice theory, this companion. More generally, we can look for ways of attaching to $T$ fairly nice theories which have some of the properties of this missing model companion.

- When the theory $T$ has a model companion, we attach to $T$ a particularly nice elementary class of structures, the models of the companion. More generally, we can look for ways of attaching to $T$ fairly nice (but non-elementary) classes of structures theories which have some of the properties of this missing elementary class.

In this section we begin to develop the first approach. We begin the development of the second approach in the next section.

How can we attach to a theory a fairly nice companion theory? Whatever we do the selection should be done in a uniform way.

6.11 DEFINITION. For a fixed language $L$, a companion operator is an assignment $T \rightarrow T^a$ between theories (of the underlying language) such that
(i) \(T\) and \(T'\) are companions

(ii) if \(T_1\) and \(T_2\) are companions, then \(T_1' = T_2'\)

(iii) \(T \cap \forall \subseteq T'\)

for all theories \(T, T_1, T_2\).

Thus a companion operator selects from each companion block a particular member. Furthermore, it ensures that the \(\forall\)-apart of the selected theory is as large as possible. Notice that because of this third condition it is not immediately obvious that there are any companion operators. Shortly we will produce a rather syntactic example of a companion operator. Before that let’s see why these gadgets are useful.

The next result follows by a simple property of companion operators, see Exercise 6.3.

6.12 LEMMA. Let \((\cdot)^a\) and \((\cdot)^b\) be a pair of companion operators (for a given language). Then we have

\[T^a \cap \forall = T^b \cap \forall\]

for each theory \(T\) (of that language).

The next result is a simple consequence of Corollary 6.3, but it deserves a higher status.

6.13 THEOREM. Let \((\cdot)^a\) be a companion operator (for some language \(L\)). If the theory \(T\) has a model companion \(T^*\), then \(T^a = T^*\).

This indicates why the notion of a companion operator is useful. If we are looking for a (possibly non-existing) model companion of a theory, then we hit that theory with one or several companion operators at our disposal. If the theory has a model companion then the hit will come up trumps.

That’s all very nice, but as yet we don’t have any example of a companion operator.

Where can we find an example of a companion operator? By Lemma 6.12, for each companion operator \((\cdot)^a\) the set \(T^a \cap \forall\) depends only on the theory \(T\) and not on the particular operator \((\cdot)^a\). We can construct this set directly, and this leads to the minimum companion operator.

6.14 DEFINITION. For a theory \(T\) an \(\forall\)-sentence \(\sigma\) is 0-tame over \(T\) if we have

\[T \cap \forall_1 \vdash (\sigma \rightarrow \alpha) \implies T \vdash \alpha\]

for each \(\forall_1\)-sentence \(\alpha\). Let \(0(T)\) be the set of all such \(\forall_2\)-sentences.

Thus the \(\forall_2\)-sentence \(\sigma\) is 0-tame over \(T\) if and only if

\[(T \cap \forall_1) \cup \{\sigma\}\]

axiomatizes a companion of \(T\). Notice also that \(T \cap \forall_2 \subseteq 0(T)\).

At first sight the following result can be a little surprising.

6.15 LEMMA. For each (consistent) theory \(T\) the set \(0(T)\) is closed under conjunction, consistent, and axiomatizes a companion of \(T\).
Proof. To begin we show that
\[ \sigma, \tau \in 0(T) \implies \sigma \land \tau \in 0(T) \]
for all \( \forall_2 \)-sentences \( \sigma, \tau \).

Consider \( \sigma, \tau \in 0(T) \). Trivially, \( \sigma \land \tau \) is a \( \forall_2 \)-sentence, and every model of
\[ (T \cap \forall_1) \cup \{ \sigma \land \tau \} \]
is a submodel of \( T \). Thus it is sufficient to show the converse, that each model of \( T \) is a submodel of this set.

Consider any \( \mathfrak{A} \in S(T) \). Using first \( \sigma \) and then \( \tau \) we we obtain
\[ \mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{C} \quad \mathfrak{B} \models \sigma \quad \mathfrak{C} \models \tau \]
for some submodels \( \mathfrak{B}, \mathfrak{C} \in S(T) \). By iteration we generate two interlacing chains
\[ B = \{ \mathfrak{B}_i \mid i < \omega \} \quad C = \{ \mathfrak{C}_i \mid i < \omega \} \]
of submodels of \( T \) such that
\[ B_i \models \sigma \quad C_i \models \tau \]
for each \( i < \omega \). Let \( \mathcal{U} \) be the common union of these two chains. Then
\[ \mathfrak{A} \subseteq \mathcal{U} \in S(T) \quad \mathcal{U} \models \sigma \land \tau \]
to give the required result.

By way of contradiction suppose \( 0(T) \) is not consistent. Thus
\[ 0(T) \vdash \text{false} \]
so that (by the first part) we have
\[ \vdash \sigma \rightarrow \text{false} \]
for some \( \sigma \in 0(T) \). But now
\[ T \cap \forall_1 \vdash \sigma \rightarrow \text{false} \]
to give
\[ T \vdash \text{false} \]
which is not so (since \( T \) is consistent).

By construction we have
\[ T \cap \forall_2 \subseteq 0(T) \]
and \( 0(T) \) depends only on \( T \cap \forall_1 \). It remains to show that \( 0(T) \) axiomatizes a companion of \( T \). We have
\[ T \cap \forall_1 \subseteq T \cap \forall_2 \subseteq 0(T) \]
so to complete the proof it suffices to show that
\[ 0(T) \vdash \alpha \iff \alpha \in T \]
for each \( \forall_1 \)-sentence \( \alpha \). To this end, consider any \( \forall_1 \)-sentence \( \alpha \) with \( 0(T) \vdash \alpha \). Since \( 0(T) \) is closed under conjunction, there is some \( \sigma \in 0(T) \) such that \( \vdash \sigma \rightarrow \alpha \). But now \( \alpha \in T \) since \( \sigma \) is 0-tame over \( T \).

Naturally, we use \( 0(T) \) as a set of axiom for a theory.
6.16 **DEFINITION.** For each theory $T$ let $T^0$ be the theory axiomatized by $0(T)$. ■

This gives us the first example of a companion operator. In fact, it is the least companion operator.

6.17 **THEOREM.** For each language $L$ the assignment $(\cdot)^0$ is a companion operator. Furthermore, $T^0 \subseteq T^a$ for each theory $T$ and each companion operator $(\cdot)^a$.

**Proof.** By Lemma 6.15, the two theories $T$ and $T^0$ are companions. By construction, $T^0$ depends only on $T \cap \forall_1$, and hence $T_1^0 = T_2^0$ for companions $T_1, T_2$. By construction, $T \cap \forall_2 \subseteq T^0$.

This shows that $(\cdot)^0$ is a companion operator.

Finally, for each companion operator $(\cdot)^a$, we have

$$T^0 \subseteq T^{0a} = T^a$$

using simple properties of such an operator. ■

As we proceed we will see several more companion operators, all of which are more interesting than $(\cdot)^0$.

**Exercises**

6.3 Let $(\cdot)^a$ and $(\cdot)^b$ be companion operators. Show that

$$T^{ab} = T^b \quad T^a \cap \forall_2 = T^b \cap \forall_2$$

hold for all theories $T$.


6.5 Show that for each companion operator $(\cdot)^a$ and each theory $T$, the companion $T^0$ is the theory axiomatized by $T^a \cap \forall_2$.

6.3 **Existentially closed structures**

When a theory $T$ has a model companion $T^*$, we have a particularly nice elementary subclass $\mathcal{M}_d(T^*)$ of $S(T)$. What can we do when $T$ does not have a model companion? We can look for a subclass of $S(T)$ which has most of the characteristic properties of $\mathcal{M}_d(T^*)$, except that it need not be elementary. There are several possible such subclasses. In this section we look at one of the simplest.

6.18 **DEFINITION.** A structure $\mathfrak{A}$ is existentially closed for a theory $T$ if $\mathfrak{A} \in S(T)$ and we have

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec_1 \mathfrak{B}$$

for each model $\mathfrak{B} \models T$.

Let $\mathcal{E}(T)$ be the class of structures which are existentially closed for $T$. ■

The definition of $\mathcal{E}(T)$ makes specific reference to models $\mathfrak{B}$ of the theory $T$. In fact, this is not necessary. The proof of the following is a simple exercise.
6.19 **Lemma.** For each theory $T$ and $\mathfrak{A} \in \mathcal{E}(T)$ we have

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec_1 \mathfrak{B}$$

for all $\mathfrak{B} \in \mathcal{S}(T)$.

This shows that $\mathcal{E}(T)$ depends only on the class $\mathcal{S}(T)$ and not the smaller class $\mathcal{M}d(T)$. In other words, $\mathcal{E}(T)$ depends only on the companion block of $T$.

6.20 **Theorem.** For a theory $T$ with a model companion $T^*$ we have $\mathcal{E}(T) = \mathcal{M}d(T^*)$.

**Proof.** Consider any $\mathfrak{A} \in \mathcal{M}d(T^*)$ and any $\mathfrak{B} \models T$ with $\mathfrak{A} \subseteq \mathfrak{B}$. Since $T$ and $T^*$ are companions, we have $\mathfrak{B} \subseteq \mathfrak{C}$ for some $\mathfrak{C} \models T^*$. But $T^*$ is model complete, so $\mathfrak{A} \prec \mathfrak{C}$, and hence $\mathfrak{A} \prec_1 \mathfrak{B}$. Thus $\mathfrak{A} \in \mathcal{E}(T)$. This shows that $\mathcal{M}d(T^*) \subseteq \mathcal{E}(T)$.

Conversely, consider any $\mathfrak{A} \in \mathcal{E}(T)$. Since $T$ and $T^*$ are companions, we have $\mathfrak{A} \subseteq \mathfrak{B}$ for some model $\mathfrak{B} \models T^*$. Then $\mathfrak{A} \prec_1 \mathfrak{B}$, so that $\mathfrak{B} \models (\forall_2) \mathfrak{A}$, and hence $\mathfrak{A} \models T^*$ (since $T^*$ is $\forall_2$-axiomatizable). Thus $\mathfrak{A} \in \mathcal{M}d(T^*)$. This shows that $\mathcal{E}(T) \subseteq \mathcal{M}d(T^*)$. ■

At this point we should show that $\mathcal{E}(T)$ is non-empty for every theory $T$. We should prove the following.

6.21 **Theorem.** For each theory $T$ the class $\mathcal{E}(T)$ is cofinal in $\mathcal{S}(T)$. In other words, for each $\mathfrak{A} \in \mathcal{S}(T)$ there is some $\mathfrak{B} \in \mathcal{E}(T)$ with $\mathfrak{A} \subseteq \mathfrak{B}$.

In order not to upset the flow, we will postpone the proof of Theorem 6.21 until the next section. Here we continue to develop the properties of $\mathcal{E}(\cdot)$. Of course, to do this we will occasionally use this existence result.

6.22 **Lemma.** For each theory $T$ we have

$$\mathfrak{A} \prec_1 \mathfrak{B} \in \mathcal{E}(T) \implies \mathfrak{A} \in \mathcal{E}(T)$$

for all structures $\mathfrak{A}, \mathfrak{B}$.

**Proof.** Consider any situation

$$\mathfrak{A} \prec_1 \mathfrak{B} \in \mathcal{E}(T) \quad \mathfrak{A} \subseteq \mathfrak{C} \models T$$

for structure $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$. We must show that $\mathfrak{A} \prec_1 \mathfrak{C}$.

We have a wedge, as to the left

$$\begin{array}{ccc}
\mathfrak{B} & \rightarrow & \mathfrak{C} \\
\downarrow f & & \downarrow g \\
\mathfrak{A} & \rightarrow & \mathfrak{A}
\end{array}$$

where $f$ is a $\prec_1$-embedding and $g$ is an embedding. (In fact, both are insertions.) By Lemma 4.7 there is a commuting square of embeddings, as to the right, where $k$ is elementary. In particular, $\mathfrak{C} \models T$, and hence $h$ is a $\prec_1$-embedding (since $\mathfrak{B} \in \mathcal{E}(T)$.)
Now consider any $\forall_1$-formula $\phi(v)$ and point $a$ of $\mathfrak{A}$ (which matches $v$). Using the various kinds of embeddings we have

$$
\mathfrak{A} \models \phi(a) \implies \mathfrak{B} \models \phi(fa) \implies \mathfrak{D} \models \phi((h \circ f)a) \implies \mathfrak{D} \models \phi((k \circ g)a) \implies \mathfrak{C} \models \phi(ga)
$$

so that

$$
\mathfrak{A} \models \phi(a) \implies \mathfrak{C} \models \phi(a)
$$

and hence $\mathfrak{A} \prec_1 \mathfrak{C}$, as required. ■

With this result we can produce an intrinsic characterization of $\mathcal{E}(\cdot)$.

**6.23 THEOREM.** For each theory $T$ the class $\mathcal{E}(T)$ is uniquely characterized by the following three properties.

(i) $\mathcal{E}(T)$ is cofinal in $\mathcal{S}(T)$.

(ii) We have

$$
\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec_1 \mathfrak{B}
$$

for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{E}(T)$.

(iii) We have

$$
\mathfrak{A} \prec_1 \mathfrak{B} \in \mathcal{E}(T) \implies \mathfrak{A} \in \mathcal{E}(T)
$$

for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{S}(T)$.

**Proof.** We must do two things. We must show that the class $\mathcal{E}(T)$ has these three properties, and we must show that $\mathcal{E}(T)$ is the only class with these three properties.

First we verify that $\mathcal{E}(T)$ has (i, ii, iii).

(i) This is the basic existence result, Theorem 6.21 (which, of course, we have not yet proved).

(ii) As in Lemma 6.19, this is a simple consequence of the definition of $\mathcal{E}(T)$.

(iii) This is just Lemma 6.22.

Secondly, let $\mathcal{E}'(T)$ be any class with properties (i', ii', iii') corresponding to (i, ii, iii). We must show that $\mathcal{E}(T) = \mathcal{E}'(T)$.

Consider any $\mathfrak{A} \in \mathcal{E}(T)$. By (i') and (i) there are structures

$$
\mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{C}
$$

where $\mathfrak{B} \in \mathcal{E}'(T)$ and $\mathfrak{C} \in \mathcal{E}(T)$. By (ii) we have $\mathfrak{A} \prec_1 \mathfrak{C}$, and hence $\mathfrak{A} \prec_1 \mathfrak{B}$. But now (iii') gives $\mathfrak{A} \in \mathcal{E}'(T)$.

This gives $\mathcal{E}(T) \subseteq \mathcal{E}'(T)$, and a symmetric argument gives the converse inclusion. ■

In general, the class $\mathcal{E}(T)$ is not elementary. When it is elementary it is the class of models of the model companion of $T$. However, even though the class need not be elementary, it still has a theory.

**6.24 DEFINITION.** For each theory $T$ let $T^e = \text{Th}(\mathcal{E}(T))$. ■
When $T$ has a model companion $T^*$, we have $\mathcal{E}(T) = \mathcal{M}d(T^*)$, and then $T^e = T^*$. In fact, this equality is a consequence of the following (whose proof is left as an exercise).

6.25 **THEOREM.** *For each language $L$ the assignment $(\cdot)^e$ is a companion operator.*

Existentially closed structures are an important tool in model theory. We will use them quite a lot, and develop some special classes of these structures.

**Exercises**

6.6  Prove Lemma 6.19.

6.7  Prove a converse to Theorem 6.20. In other words, show that for each theory $T$, if the class $\mathcal{E}(T)$ is elementary, then $T$ has a model companion.

6.8  Show that for each theory $T$ the class $\mathcal{E}(T)$ is closed under unions of directed systems.

6.9  Let $T$ be an arbitrary theory.

(a) Show that

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b) \implies (\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

holds for each $\mathfrak{A} \in \mathcal{E}(T), \mathfrak{B} \in \mathcal{S}(T)$, and matching points $a, b$ of these structures.

(b) Show that

$$\mathfrak{A} \equiv (\exists_1) \mathfrak{B} \implies \mathfrak{A} \equiv_2 \mathfrak{B}$$

holds for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{E}(T)$.

6.10  Prove Theorem 6.25.

6.11  Let $T$ be the theory of graphs (as in Section 5.3). Show that each member of $\mathcal{E}(T)$ is random.

6.12  Let $T$ be the theory of commutative rings (as in Example 6.6), and let

$$\text{NNil}(v) = \{(v^n \neq 0) \mid n \in \mathbb{N}\}$$

to obtain a quantifier-free type.

Show that

$$\mathcal{E}(T) \models (\forall v)[\bigwedge \text{NNil}(v) \leftrightarrow \text{nnil}(v)]$$

(where this infinite conjunction has the obvious semantics).

Does this show that

$$T^e \cup \text{NNil}(v) \vdash \text{nnil}(v)$$

holds?
6.4 Existence of existentially closed structures

In this section we prove Theorem 6.21. Thus we show that for each theory \( T \) the class \( \mathcal{E}(T) \) is cofinal in \( \mathcal{S}(T) \). To do that we first obtain a characterization of a structure being existentially closed, namely that its \( \forall_1 \)-behaviour is controlled by its \( \exists_1 \)-behaviour. This is made precise as follows.

6.26 THEOREM. Let \( T \) be a theory and let \( \mathfrak{A} \in \mathcal{S}(T) \). The following are equivalent.

(i) \( \mathfrak{A} \in \mathcal{E}(T) \)

(ii) For each \( \forall_1 \)-formula \( \phi(v) \) and point \( a \) of \( \mathfrak{A} \) with \( \mathfrak{A} \models \phi(a) \), we have

\[
T \cup \text{Diag}(\mathfrak{A}, a) \vdash \phi(a)
\]

where \( a \) is a full enumeration of \( \mathfrak{A} \).

(iii) For each \( \forall_1 \)-formula \( \phi(v) \) and point \( a \) of \( \mathfrak{A} \) with \( \mathfrak{A} \models \phi(a) \), we have

\[
\mathfrak{A} \models \theta(a) \quad T \vdash \theta \rightarrow \phi
\]

for some \( \exists_1 \)-formula \( \theta(v) \).

Proof. (i)\( \Rightarrow \)(ii). Assuming (i) and \( \mathfrak{A} \models \phi(a) \), consider any model of \( T \cup \text{Diag}(\mathfrak{A}, a) \). In other words, consider a structure \( \mathfrak{B} \) with \( \mathfrak{A} \subseteq \mathfrak{B} \models T \). Then (i) gives \( \mathfrak{A} \prec_1 \mathfrak{B} \) so that \( \mathfrak{B} \models \phi(a) \) for the required result.

(ii)\( \Rightarrow \)(iii). Assuming (ii) consider any \( \forall_1 \)-formula \( \phi(v) \) where \( \mathfrak{A} \models \phi(a) \) for some point \( a \) of \( \mathfrak{A} \). By (ii) we have

\[
T \cup \text{Diag}(\mathfrak{A}, a) \vdash \phi(a)
\]

and hence

\[
T \vdash \delta(w, v) \rightarrow \phi(v)
\]

for some quantifier-free formula \( \delta(w, v) \) such that \( \mathfrak{A} \models \delta(b, a) \) for some point \( b \) of \( \mathfrak{A} \). Let

\[
\theta(v) = (\exists w)\delta(w, v)
\]

to produce the required \( \exists_1 \)-formula.

(iii)\( \Rightarrow \)(i). Assuming (iii) consider structure a \( \mathfrak{B} \) with \( \mathfrak{A} \subseteq \mathfrak{B} \models T \). Suppose \( \mathfrak{A} \models \phi(a) \) where \( \phi(v) \in \forall_1 \) and \( a \) is a point of \( \mathfrak{A} \). We require \( \mathfrak{B} \models \phi(a) \). But, using the formula \( \theta(v) \) given by (iii), we have

\[
\mathfrak{A} \models \theta(a)
\]

so that

\[
\mathfrak{B} \models \theta(a)
\]

since \( \theta \) is \( \exists_1 \), and hence

\[
\mathfrak{B} \models \phi(a)
\]

since \( \mathfrak{B} \models T \).

We use this controlling idea to embed each submodel of a theory into an existentially closed structure of that theory. As with many of the constructions we use this one
comes in two parts. A 1-step construction which partly solves the problem. Then an accumulation construction, in which the 1-step construction is iterated to closure, and solves the problem in full.

You should read the statement of the next result with some care, and observe the different roles played by \(A\) and \(A'\).

6.27 LEMMA. (The 1-step construction) Let \(T\) be a theory. For each \(A \in S(T)\), there is a model \(A'\) of \(T\) with \(A \subseteq A'\), and such that for each \(\forall_1\)-formula \(\phi(v)\) and point \(a\) of \(A\) with \(A' \models \phi(a)\) we have

\[
A' \models \theta(a) \quad T \vdash \theta \rightarrow \phi
\]

for some \(\exists_1\)-formula \(\theta(v)\).

Proof. Given the structure \(A \in S(T)\), we extend the underlying language \(L\) to \(L(a)\) by adding a parameter to name each member of \(A\). Thus \(\Delta = \text{Diag}(A, a)\) is a set of quantifier-free \(L(a)\)-sentences. Consider those sets \(\Psi\) of \(\exists_1\)-sentences of \(L(a)\) such that both

\[
\Delta \subseteq \Psi \quad T \cup \Psi \text{ is consistent}
\]

hold. In particular, \(\Delta\) is one such set. This family of sets is partially ordered by inclusion, and is closed under unions of directed families.

By Zorn’s Lemma, there is a maximal such set \(\Psi\).

Consider any model of \(T\) and of this maximal set. This gives a structure \(A \subseteq A' \models T\) such that \((A', a) \models \Psi\). We show that this structure has the partial control property.

Consider any \(\forall_1\)-formula \(\phi(v)\) (of the parent language \(L\)), consider any point \(a\) from \(A\) which matches the free variables \(v\), and suppose \(A' \models \phi(a)\). To produce the controlling \(\exists_1\)-formula \(\theta(v)\) we look at the set

\[
T \cup \Psi \cup \{\neg \phi(a)\}
\]

of sentences in \(L(a)\).

If this set is consistent, then (since \(\neg \phi \in \exists_1\)) the maximality of \(\Psi\) gives \(\neg \phi(a) \in \Psi\), and hence \(A' \models \neg \phi(a)\), which is not so. Thus the set is inconsistent and hence (by compactness)

\[
T \cup \Phi \vdash \phi(a)
\]

for some finite part \(\Phi\) of \(\Psi\). The maximality of \(\Psi\) ensures that it is closed under conjunction, and hence we can replace \(\Phi\) be a singleton. Thus, moving to the language \(L\) we have

\[
T \vdash \psi(w, v) \rightarrow \phi(v)
\]

for some formula \(\psi \in \exists_1\) where \(A' \models \psi(b, a)\) for some point \(b\) of \(A'\) taken from \(A\). The formula \(\theta(v) = (\exists w)\psi(w, v)\) does the required job.

The structure \(A'\) does its best to be an existentially closed extension of \(A\), but it can only deal with points from \(A\). To get round this we repeat the trick over and over again.

6.28 THEOREM. (The accumulation construction) Let \(T\) be a theory. For each \(A \in S(T)\) there is some \(A^+ \in E(T)\) with \(A \subseteq A^+\).
Proof. Given a structure $A \in S(T)$ we may iterate the use of Lemma 6.27 to produce an $\omega$-chain

$$A = A_0 \subseteq A_1 \subseteq \cdots \subseteq A_i \subseteq \cdots$$

(i < $\omega$)

of submodels of $T$. Thus

$$A_0 = A, \quad A_{i+1} = A'_i$$

for each $i < \omega$. Here $(\cdot)'$ indicates a use of the construction of Lemma 6.27. In particular, each structure $A_{i+1}$ can be taken to be a model of $T$, but that is not particularly relevant.

Let

$$A^+ = \bigcup \{ A_i \mid i < \omega \}$$

so that $A \subseteq A^+ \in S(T)$ (and, in fact, $A' \models T \cap \forall_2$). We show that $A^+ \in E(T)$.

Consider any structure $B$ with $A^+ \subseteq B \models T$, and any $\forall_1$-formula $\phi(v)$ with $A^+ \models \phi(a)$ for some point $a$ of $A^+$. We require $B \models \phi(a)$.

Since $a$ has just finitely many elements, there is some $i < \omega$ such that $a$ comes from $A_i$. We have

$$A'_i = A_{i+1} \models \phi(a)$$

since $\phi$ is $\forall_1$. Thus, by the construction of Lemma 6.27, we have

$$A_{i+1} \models \theta(a) \quad \text{for some } \exists_1\text{-formula } \theta(v).$$

This gives $B \models \theta(a)$, and hence $B \models \phi(a)$ since $B \models T$, for the required result. $\blacksquare$

You should make sure you understand this proof. It is a rather rudimentary version of a saturation process. Later [**Say where**] we will refine the construction, and there Zorn’s Lemma will not be available.

In this and the previous section we have used the 1-embedding relation $\prec_1$ to extract certain properties of theories. You may wonder what happens if we try to do the same kind of thing using the full elementary embedding relation. Exercises 6.13 - 6.15 indicates some of the things that can happen.

Exercises

6.13 By analogy with Theorem 6.23 show that for each theory $T$ there is at most on class $G(T)$ with the following three properties.

(i) $G(T)$ is cofinal in $S(T)$.

(ii) We have

$$A \subseteq B \implies A \prec B$$

for all $A, B \in G(T)$.

(iii) We have

$$A \prec B \in G(T) \implies A \in G(T)$$

for all $A, B \in S(T)$.
6.14 For an arbitrary theory $T$ suppose $\mathcal{G}(T)$ is a class with the properties (i, ii, ii) of Exercises 6.13

(a) Show that $\mathcal{G}(T) \subseteq \mathcal{E}(T)$.
(b) Show that $\mathcal{G}(T)$ is closed under unions of directed systems.
(c) Show that

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b) \implies (\mathfrak{A}, a) \equiv (\mathfrak{B}, b)$$

for each $\mathfrak{A}, \mathfrak{B} \in \mathcal{G}(T)$ where $a$ and $b$ are a compatible pair of points.

For the following exercise you may assume that for the language under consideration the class $\mathcal{G}(T)$ exists for each theory $T$ in that language. This is true and will be proved in [**Later**]

6.15 (a) Show that setting $T^g = Th(\mathcal{G}(T))$ for each theory $T$ (in the language under consideration) produces a companion operator.

(b) Show that a theory $T$ has a model companion precisely when $\mathcal{G}(T)$ is an elementary class, in which case $T^g$ is the model companion.

6.5 The use of types

In this section we gather a bit more information about existentially closed structures. At the moment this is not too important, but the methods used will be refined and developed later.

We have used the notion of a type a few times already (sometimes without specifically mentioning the notion). However, this is the first place where we make serious use of types. So let’s pretend we have never seen the notion before.

6.29 DEFINITION. For a given language a type is a set $\Gamma$ of formulas such that

$$\partial \Gamma = \bigcup \{ \partial \phi \mid \phi \in \Gamma \}$$

is finite.

The crucial restriction is that $\partial \Gamma$ is finite. We often write $\Gamma(v)$ to indicate that $\Gamma$ is a type with $v$ as its batch of free variables. In this section we are concerned with $\exists_1$-types, those types which contain only $\exists_1$-formulas. In passing, we also meet some $\forall_1$-types.

Each point $a$ of a structure $\mathfrak{A}$ gives us several types.

6.30 DEFINITION. Let $\mathfrak{A}$ be a structure, let $a$ be a point of $\mathfrak{A}$, and let $v$ be a batch of variables matching $a$.

The set $\Gamma(v)$ of all those formulas $\phi(v)$ with

$$\mathfrak{A} \models \phi(a)$$

is full type of $a$ in $\mathfrak{A}$.

The sets

$$\Sigma(v) = \Gamma(v) \cap \exists_1 \qquad \Pi(v) = \Gamma(v) \cap \forall_1$$

are, respectively, the $\exists_1$-type and $\forall_1$-type of $a$ in $\mathfrak{A}$.
The type $\Gamma(v)$ is more or less $\text{Th}(\mathfrak{A}, a)$ but with the constants $a$ released into the community to become variables. In the same way $\Sigma(v)$ and $\forall(v)$ are more or less $\text{Th}(\mathfrak{A}, a) \cap \exists_1 \text{Th}(\mathfrak{A}, a) \cap \forall_1$ respectively. There are various obvious refinements of these notions but we don’t need then just yet.

Each point of a structure gives us several types. There is a converse question. When can a type be obtain in this way, perhaps by some point in a structure of some special kind?

6.31 DEFINITION. Let $\Gamma(v)$ be a type (perhaps of restricted quantifier complexity), and let $\mathfrak{A}$ be a structure.

(r) We say $\mathfrak{A}$ realizes $\Gamma(v)$ if $\mathfrak{A} \models \Gamma(a)$ for some point $a$ of $\mathfrak{A}$ (matching the batch $v$).

(o) We say $\mathfrak{A}$ omits $\Gamma(v)$ if it is not realized in $\mathfrak{A}$. ■

These two notions become more and more important as model theory develops. The main purpose of this section is to introduce these ideas.

6.32 DEFINITION. Let $T$ be a theory.

A maximal-$\exists_1$ type over $T$ is an $\exists_1$-type $\Sigma(v)$ which is maximally consistent with $T$, that is $T \cup \Sigma(v)$ is consistent and $T \cup \Sigma \cup \{\theta(v)\}$ is consistent $\implies \theta \in \Sigma$ for each $\exists_1$-formula $\theta$.

For each batch $v$ of variable let $\exists_1^\max(T, v)$ be the set of types $\Sigma(v)$, in the batch $v$, which are maximal-$\exists_1$ over $T$. ■

We often vary the phrase ‘maximal-$\exists_1$ type over $T$’ and refer to a $\exists_1$-type which is maximally consistent with $T$, or some similar phrase.

What has this notion got to do with existential closedness?

6.33 LEMMA. Let $T$ be a theory. For each $\exists_1$-type $\Sigma(v)$ the following are equivalent.

(i) The type $\Sigma(v)$ is $\exists_1$-maximal over $T$.

(ii) The type $\Sigma(v)$ is the $\exists_1$-type of some point $a$ of some $\mathfrak{A} \in \mathcal{E}(T)$.

Proof. $(i) \Rightarrow (ii)$. Suppose $\Sigma(v)$ is a maximal-$\exists_1$ type over $T$. Since $\Sigma$ is consistent with $T$, it is realized in some model $\mathfrak{B}$ of $T$. Consider any $\mathfrak{A} \in \mathcal{E}(T)$ with $\mathfrak{B} \subseteq \mathfrak{A}$. The $\exists_1$-type is realized in $\mathfrak{A}$ by some point $a$, say. We show that $\Sigma$ is the $\exists_1$-type of $a$ in $\mathfrak{A}$.

Consider any $\exists_1$-formula $\theta(v)$ such that $\mathfrak{A} \models \theta(a)$. Then $T \cup \Sigma \cup \{\theta(v)\}$ is consistent, and hence $(i)$ gives $\theta \in \Sigma$, as required.

$(ii) \Rightarrow (i)$. Consider any $\mathfrak{A} \in \mathcal{E}(T)$ and point $a$ of $\mathfrak{A}$, and suppose $\Sigma(v)$ is the $\exists_1$-type of $a$ in $\mathfrak{A}$. In particular, $\Sigma$ is consistent with $T$. Consider any $\exists_1$-formula $\theta(v)$ such that $T \cup \Sigma \cup \{\theta(v)\}$
is consistent. There is some model $\mathcal{B}$ of $T$ and some point $b$ of $\mathcal{B}$ such that

$$\mathcal{B} \models \Sigma(b) \quad \mathcal{B} \models \theta(b)$$

hold. The first of these gives

$$(\mathcal{A}, a) \equiv (\exists_1) (\mathcal{B}, b)$$

so that

$$(\mathcal{A}, a) \equiv_1 (\mathcal{B}, b)$$

(since $\mathcal{A} \in \mathcal{E}(T)$). But now $\mathcal{A} \models \theta(a)$, which leads to the required result. ■

Using a variant of this proof we obtain the following characterization.

**6.34 LEMMA.** Let $T$ be a theory. For each $\mathcal{A} \in \mathcal{S}(T)$ the following are equivalent.

(i) $\mathcal{A} \in \mathcal{E}(T)$.

(ii) For each point $a$ of $\mathcal{A}$, the $\exists_1$-type of $a$ in $\mathcal{A}$ is maximally consistent with $T$.

This indicates that to construct an existentially closed structure we must somehow maximize the existential types involved. That is what is going on in the proofs of Lemma 6.27 and Theorem 6.28. The maximization is achieved by a use of Zorn’s Lemma and the Accumulation process. Later, in Chapter [**saturation**] we will refine this idea to produce a saturation construction.

We also develop the omitting types process. As an introduction to this look at Exercise 6.18.

**Exercises**

6.16 Prove Lemma 6.34.

6.17 For an arbitrary theory $T$ suppose $\mathcal{G}(T)$ is a class with the properties $(i, ii, iii)$ of Exercises 6.13

Show that for each $\mathcal{A} \in \mathcal{S}(T)$ we have $\mathcal{A} \in \mathcal{G}(T)$ precisely when each formula $\phi(v)$ and point $a$ of $\mathcal{A}$ with $\mathcal{A} \models \phi(a)$, there is an $\exists_1$-type $\Sigma(v)$ such that

$$\mathcal{A} \models \Sigma(a) \quad \mathcal{G}(T) \models (\forall v)[\bigwedge \Sigma(v) \rightarrow \phi(v)]$$

where this last modelling relation is handled in the obvious way.

6.18 Let $T$ be a theory and let $\phi(v)$ be a formula in the indicated batch of variables. Let $\Sigma(v)$ be the set of $\exists_1$-formulas $\theta(v)$ such that both

$$T \cup \{\theta\} \text{ is consistent} \quad T \vdash \theta \rightarrow \phi$$

hold. Let

$$\Omega(T, \phi) = \{\phi\} \cup \{\neg \theta \mid \theta \in \Sigma\}$$

(so that $\Omega(T, \phi)$ is an $\forall_1$-type when $\phi$ is a $\forall_1$-formula).

Show that for a structure $\mathcal{A} \in \mathcal{S}(T)$, we have $\mathcal{A} \in \mathcal{E}(T)$ precisely when for each $\forall_1$-formula $\phi$ the structure $\mathcal{A}$ omits $\Omega(T, \phi)$. 
In this chapter we look at two kinds of structures which, in a sense, are at the opposite ends of the spectrum of submodels of a theory. They are the ‘small’ and the ‘large’ submodels, but not in the sense of size (cardinality). The ‘small’ ones realize as few types as possible. Formally, these are the atomic structures in an appropriate sense. The nature of the embeddings involved must be taken into account, so there are various flavours of atomic structures. The ‘large’ structures are the ones that realize as many types as possible. Formally, these are the appropriately saturated structures, where again the nature of the embeddings involved must be taken into account.

Here we look at two flavours of such structures. One version, (0), uses embeddings, and the other, \((\omega)\), uses elementary embedding. There are also many versions \((n)\) in between but, on the whole, these are not so interesting. Once we understand the 0-version, it is easy to generate the \(n\)-version.

On the whole, we will concentrate on the 0-version. This is more complicated of the two. We will merely sketch the details of the \(\omega\)-version.

It may seem more natural to deal with the small structures first. However, for at least two reasons, it is easier to deal with the large structures first.

### 7.1 Existentially universal structures

In this section we begin to look at structures that are large in the sense that each such structure contains as many different kinds of points as is compatible with its environment. These are saturated structures. There are many variants of this notion, and we concentrate on just one of them. We deal with what could be termed the \(\aleph_0\)-(0)-version. Later, in \[[**section on saturated structures**\]], we look at other variants of this notion.

We have caught a glimpse of the idea of saturation all ready when we looked at existentially closed structures. However, there the saturation is so feeble that we didn’t bother to describe it in this way.

Saturation is concerned with the realization of types. A structure is saturated if it realizes as many types (of a predetermined kind) as possible.

So far most types that we have seen have been pure, that is they have been sets of formulas of the underlying language. We now begin to use types which may contain parameters from some selected structure. In other words, we use types in various enrichments of the underlying language.

Consider a structure \(\mathfrak{A}\) for some language \(L\). Let \(a\) be a point of \(\mathfrak{A}\). This gives a finite enrichment \(L(a)\) of \(L\). Let \(\Phi(a, v)\) be a type in this enriched language. Thus we have a pure type \(\Phi(u, v)\) in a larger list of variables, and some of these are instantiated by the selected parameters. In this section we are concerned mostly with \(\exists_1\)-types in enriched languages.

### 7.1 DEFINITION

Let \(L\) be a language and let \(\mathfrak{A}\) be a \(L\)-structure.
A type over \( A \) is a type in a language \( L(a) \) formed from \( L \) by adding a list \( a \) of names for finitely many elements of \( A \).

An \( \exists_1 \)-type over \( A \) is a type over \( A \) which consists of \( \exists_1 \)-formulas of the enriched language.

How can we find examples of such \( \exists_1 \)-types over a structure \( A \)?

Fix a point \( a \) of \( A \). These will be the parameters in the types. Consider also another point \( b \) of \( A \) (which may overlap \( a \)). Thus we have an extended point \( a \downarrow b \), but it is convenient to keep the two parts separate. Let \( \Theta(u, v) \) be the \( \exists_1 \)-type in \( A \) of this extended point. Thus \( \Theta(u, v) \) consists of all \( \exists_1 \)-formulas \( \theta(u, v) \) such that \( A \models \theta(a, b) \).

The set 
\[
\Theta(a, v)
\]
is a \( \exists_1 \)-types over \( A \). In fact, this type is realized in \( A \) since
\[
A \models \Theta(a, b)
\]
holds for some point \( b \). We may write
\[
A \models (\exists v) \land \Theta(a, v)
\]
to indicate this without naming the point \( b \).

There may be other types which are not realized in \( A \). Consider any structure \( A \subseteq B \). With the fixed point \( a \) from \( A \), and any point \( b \) from \( B \), let \( \Theta(a, v) \) be the \( \exists_1 \)-type of \( b \) in \( B \). Thus \( B \) realizes this type
\[
B \models (\exists v) \land \Theta(a, v)
\]
but it may be omitted by \( A \).

7.2 DEFINITION. Let \( T \) be a theory, and let \( A \in S(T) \). A \( \exists_1 \)-type \( \Theta(a, v) \) over \( A \) is \( T \)-consistent over \( A \) if there is some \( A \subseteq B \models T \) in which \( \Theta(a, v) \) is realized.

Since we are dealing with \( \exists_1 \)-types, the restriction \( B \in Md(T) \) on the realizing structure can be weakened to \( B \in S(T) \).

With these preliminaries we can make precise the idea that a structure contains as many different kinds of points as possible.

7.3 DEFINITION. A structure \( A \) is existence universally for a theory \( T \) if \( A \in S(T) \) and if \( A \) realizes each \( \exists_1 \)-type over \( A \) which is \( T \)-consistent over \( A \).

Let \( U(T) \) be the class of structures which are existentially universal for \( T \).

You should compare this notion with that of an existentially closed structure as given by Definition 6.18. Although the two definitions are phrased rather differently, the two notions are quite similar. The crucial difference is that an existentially closed structure is concerned only with the realization of a single \( \exists_1 \)-formula, whereas an existentially universal structure is concerned with the realization of a whole \( \exists_1 \)-type. In particular, if we consider the case where the \( \exists_1 \)-type is a single formula, then we obtain the following.

7.4 LEMMA. For each theory \( T \) we have \( U(T) \subseteq E(T) \).
Eventually we will show that $\mathcal{U}(T)$ is cofinal in $\mathcal{S}(T)$. The proof is a refinement of that of Theorem 6.28 and quite intricate. Thus we postpone the proof until Section 7.3.

We often write $\mathfrak{M}, \mathfrak{N}, \ldots$ to indicate that we are dealing with a existentially universal structures.

Existentially universal structures have rather strong back-and-forth properties. Here is the crucial result.

**7.5 LEMMA.** Let $T$ be a theory and consider a situation

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{M}, b)$$

where $\mathfrak{A} \in \mathcal{S}(T)$, $\mathfrak{M} \in \mathcal{U}(T)$, and $a, b$ are matching points. Then for each element $x$ of $\mathfrak{A}$ we have

$$(\mathfrak{A}, a \leadsto x) \equiv (\exists_1) (\mathfrak{M}, b \leadsto y)$$

for at least one element $y$ of $\mathfrak{M}$.

**Proof.** Let $\Theta(a, v)$ be the $\exists_1$-type of $x$ in $(\mathfrak{A}, a)$. Thus

$$\Theta(a, x) = Th(\mathfrak{A}, a, x) \cap \exists_1$$

with $\mathfrak{A} \models \Theta(a, x)$. Consider the corresponding type $\Theta(b, v)$ over $\mathfrak{M}$ obtained by replacing the $\mathfrak{A}$-point $a$ be the $\mathfrak{M}$-point $b$.

Since

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{M}, b)$$

the type $\Theta(b, v)$ is finitely satisfiable in $(\mathfrak{M}, b)$, and hence finitely satisfiable in some elementary extension of $(\mathfrak{M}, b)$. Since $\mathfrak{M} \in \mathcal{U}(T)$ we have

$$\mathfrak{M} \models \Theta(b, y)$$

for at least one element $y$ of $\mathfrak{M}$, as required. $\blacksquare$

This simple result has several consequence.

By iterating the construction we obtain the following.

**7.6 THEOREM.** For each theory $T$, each $\mathfrak{M} \in \mathcal{U}(T)$, each countable $\mathfrak{A} \in \mathcal{S}(T)$, we have

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{M}, b) \implies (\mathfrak{A}, a) \text{ is embeddable in } (\mathfrak{M}, b)$$

for each pair $a, b$ of matching points from $\mathfrak{A}, \mathfrak{M}$.

We say a structure $\mathfrak{M}$ is universal for countable submodels of a theory if each such structure is embeddable in $\mathfrak{M}$. The two uses of ‘universal’ here is no coincidence.

**7.7 COROLLARY.** If the theory $T$ has JEP, the each $\mathfrak{M} \in \mathcal{U}(T)$ is universal for countable submodels of $T$.

Of course, Lemma 7.5 enables us to use the back and forth technique.
Let $\mathcal{T}$ be a theory. For each $\mathcal{M}, \mathcal{N} \in \mathcal{U}(\mathcal{T})$ we have
\[(\mathcal{M}, a) \equiv (\exists_1)(\mathcal{N}, b) \implies (\mathcal{M}, b) \equiv_p (\mathcal{N}, b)\]
for all matching points $a$ from $\mathcal{M}$ and $b$ from $\mathcal{N}$.

If $\mathcal{T}$ has JEP then we have
\[(\mathcal{M}, a) \equiv_0 (\mathcal{N}, b) \implies (\mathcal{M}, b) \equiv_p (\mathcal{N}, b)\]
for each $\mathcal{M}, \mathcal{N} \in \mathcal{U}(\mathcal{T})$ and matching points $a$ and $b$.

Proof. Since $\mathcal{M}, \mathcal{N}$ are e. c. we have
\[(\mathcal{M}, a) \equiv (\exists_1)(\mathcal{N}, b) \implies (\mathcal{M}, a) \equiv_2 (\mathcal{N}, b)\]
and Lemma 7.5 enables us to set up a back-and-forth system.

When $\mathcal{T}$ has JEP we have
\[(\mathcal{M}, a) \equiv_0 (\mathcal{N}, b) \implies (\mathcal{M}, a) \equiv_2 (\mathcal{N}, b)\]
For each $\mathcal{M}, \mathcal{N} \in \mathcal{U}(\mathcal{T})$.

This result ensures that each existentially universal structure has quite a lot of homogeneity.

Let $\mathcal{T}$ be a theory, and consider any $\mathcal{M} \in \mathcal{U}(\mathcal{T})$.

We have
\[(\mathcal{M}, a) \equiv_1 (\mathcal{M}, b) \implies (\mathcal{M}, a) \equiv_p (\mathcal{M}, b)\]
for each pair $a, b$ of matching points of $\mathcal{M}$.

If $\mathcal{T}$ has JEP then we have
\[(\mathcal{M}, a) \equiv_0 (\mathcal{M}, b) \implies (\mathcal{M}, a) \equiv_p (\mathcal{M}, b)\]
for each pair $a, b$ of matching points of $\mathcal{M}$.

We know that embeddings between existentially closed structures have reasonable strong preservation properties. Theorem 7.8 shows these are much stronger for existentially universal structures.

For each theory $\mathcal{T}$ we have
\[\mathcal{M} \subseteq \mathcal{N} \implies \mathcal{M} \prec \mathcal{N}\]
for each $\mathcal{M}, \mathcal{N} \in \mathcal{U}(\mathcal{T})$.

Notice how the back-and-forth technique and existentially universal structures seemed to be made for each other. Perhaps there is a love story waiting to be written.
7.2. A companion operator

Exercises

In the following exercises you may assume that for each theory $T$ the class $U(T)$ is cofinal in $S(T)$.

7.1 For an arbitrary theory $T$ and $\mathfrak{A} \in S(T)$, show that $\mathfrak{A} \in E(T)$ precisely when $\mathfrak{A} \prec M$ for some $M \in U(T)$.

7.2 For an arbitrary theory $T$ consider $\mathfrak{A} \in S(T)$. Suppose that for each $n \in \mathbb{N}$ there is some $\mathfrak{N}_n \in U(T)$ with $\mathfrak{A} \prec_{n+1} \mathfrak{N}_n$. Show that

$$\mathfrak{A} \subseteq \mathfrak{M} \implies \mathfrak{A} \prec \mathfrak{M}$$

for each $\mathfrak{M} \in U(T)$.

7.3 Let $T$ be an arbitrary theory. For each $n \in \mathbb{N}$ let $E_n(T)$ be the class of those structures $\mathfrak{A} \in S(T)$ with $\mathfrak{A} \prec_{n+1} \mathfrak{N}$ for some $\mathfrak{N} \in U(T)$.

(a) Show that $E_0(T) = E(T)$.

(b) Show that $E_{n+1}(T) \subseteq E_n(T)$ for each $n \in \mathbb{N}$.

(c) Show that for each $\mathfrak{A} \in (T)$ the following

(i) $\mathfrak{A} \in \bigcap\{E_n(T) \mid n \in \mathbb{N}\}$

(ii) We have

$$\mathfrak{A} \subseteq \mathfrak{M} \implies \mathfrak{A} \prec \mathfrak{M}$$

for each $\mathfrak{M} \in U(T)$.

(iii) We have $\mathfrak{A} \prec \mathfrak{N}$ for some $\mathfrak{N} \in U(T)$.

are equivalent.

7.2 A companion operator

The class $U(T)$ seems to have some nice properties. In particular, Corollary 7.10 suggest that $U(T)$ is something like the class of models of a model complete theory. Or it would be if we proved the following.

7.11 THEOREM. For each theory $T$ the class $U(T)$ is cofinal in $S(T)$.

We won’t prove this result until the next section where the proof will also proved a bit more information about existentially universal structures. However, we will make use of Theorem 7.11 in this section.

As mentioned, Corollary 7.10 in conjunction with Theorem 7.11 does suggests that $U(T)$ is getting close to a model companion for the parent theory. However, $U(T)$ need not be elementary even when $T$ does have a model companion. The class $U(T)$ needs to be filled out a bit. This is hinted at in Exercises 7.1, 7.2, and 7.3. In this section we make this precise by setting down the precise details.
7.12 **DEFINITION.** For each theory $T$ let $\mathcal{G}(T)$ be given by
\[ \mathfrak{A} \in \mathcal{G}(T) \iff \mathfrak{A} \prec \mathfrak{M} \text{ for some } \mathfrak{M} \in \mathcal{U}(T) \]
to produce a subclass of $\mathcal{S}(T)$.

We have
\[ \mathcal{U}(T) \subseteq \mathcal{G}(T) \subseteq \mathcal{E}(T) \]
where the right hand inclusion holds since each existentially universal structure is existentially closed.

The class $\mathcal{G}(T)$ is the one that appeared in the three Exercises 6.13, 6.14, 6.17, and we have the following analogue of Theorem 6.23.

7.13 **THEOREM.** For each theory $T$ the class $\mathcal{G}(T)$ is uniquely characterized by the following three properties.

(i) $\mathcal{G}(T)$ is cofinal in $\mathcal{S}(T)$.

(ii) We have
\[ \mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec \mathfrak{B} \]
for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{G}(T)$.

(iii) We have
\[ \mathfrak{A} \prec \mathfrak{B} \in \mathcal{G}(T) \implies \mathfrak{A} \in \mathcal{G}(T) \]
for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{S}(T)$.

**Proof.** By Exercise 6.13 there is at most one such class $\mathcal{G}(T)$, so it suffices to show the class given by Definition 7.12 has these three properties.

(i) This is an immediate consequence of the existence result Theorem 7.11 (which, of course, we have not yet proved).

(ii) We can actually obtain something slightly more general, and this is worth stating separately, as Theorem 7.14.

(iii) This is an immediate consequence of the definition of $\mathcal{G}(T)$.

Within this proof we promised something slightly better than property (ii). Here it is.

7.14 **THEOREM.** Let $T$ be a theory. For each $\mathfrak{A}, \mathfrak{B} \in \mathcal{G}(T)$ we have
\[ (\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b) \implies (\mathfrak{A}, b) \equiv (\mathfrak{B}, b) \]
for all matching points $a$ from $\mathfrak{A}$ and $b$ from $\mathfrak{B}$.

If $T$ has JEP then we have
\[ (\mathfrak{A}, a) \equiv_0 (\mathfrak{B}, b) \implies (\mathfrak{A}, b) \equiv (\mathfrak{B}, b) \]
for each $\mathfrak{M}, \mathfrak{N} \in \mathcal{U}(T)$ and matching points $a$ and $b$. 
7.2. A companion operator

Proof. Suppose

\((A, a) \equiv (\exists_1) (B, b)\)

where \(A, B \in \mathcal{G}(T)\). Consider \(A \prec M \in \mathcal{U}(T)\) and \(B \prec N \in \mathcal{U}(T)\). We have

\((M, a) \equiv (\exists_1) (N, b)\)

and hence

\((M, a) \equiv (N, b)\)

by Theorem 7.8, to give

\((A, a) \equiv (B, b)\)

as required.

Similar reasoning gives the second part using the second part of Theorem 7.8. ■

This extension property leads to a characterization of the members of \(\mathcal{G}(T)\). This result should be compared with Theorem 6.26.

7.15 THEOREM. Let \(T\) be a theory and let \(A \in \mathcal{S}(T)\). The following are equivalent.

(i) \(A \in \mathcal{G}(T)\)

(ii) For each formula \(\phi(v)\) and point \(a\) of \(A\) with \(A \models \phi(a)\) we have

\[A \models \Theta(a) \quad \mathcal{G}(T) \models (\forall v)[\exists \Theta \rightarrow \phi]\]

for some \(\exists_1\)-type \(\Theta(v)\).

Proof. (i)⇒(ii). Suppose \(A \in \mathcal{G}(T)\) and \(A \models \phi(a)\) for some formula \(\phi(v)\) and point \(a\) of \(A\). Let \(\Theta(v)\) be the \(\exists_1\)-type of \(a\) in \(A\). It suffices to show that

\[\mathcal{G}(T) \models (\forall v)[\exists \Theta \rightarrow \phi]\]

holds.

Consider any \(B \in \mathcal{G}(T)\) and any point \(b\) of \(B\) with \(B \models \Theta(b)\). We have

\((A, a) \equiv (\exists_1) (B, b)\)

and hence

\((A, a) \equiv (B, b)\)

by Theorem 7.14, to give \(B \models \phi(b)\), as required.

(ii)⇒(i). Suppose \(A\) has (ii), and consider any \(M \in \mathcal{U}(T)\) with \(A \subseteq M\). Then \(A \prec M\), so that \(A \in \mathcal{G}(T)\). ■

Of course, we have a substitute model companion.

7.16 DEFINITION. For each theory \(T\) let \(T^g = \text{Th}(\mathcal{G}(T))\).

The associated result is proved in the same way as Theorem 6.25.

7.17 THEOREM. For each language \(L\) the assignment \((\cdot)^g\) is a companion operator.

The members of the class \(\mathcal{G}(T)\) were originally obtained in a different way. They were constructed using the method of infinite forcing, and they are the generic structures for that technique. This is where the ‘\(\mathcal{G}(\cdot)\)’ and the ‘\((\cdot)^g\)’ comes from. However, although various kinds of forcing provide some useful techniques, that particular use was rather silly. It was soon forgotten and replaced by a saturation technique, as we describe in the next section.
7.4 Show that for each theory $T$, the class $\mathcal{G}(T)$ is elementary if and only if $T$ has a model companion, in which case $T'$ is the model companion.

7.5 Show that for each theory $T$, the class $\mathcal{G}(T)$ is closed under unions of directed systems.

7.3 Existence of existentially universal structures

In this section we prove Theorem 7.11. Thus for an arbitrary theory we show that each $A \in S(T)$ can be embedded in some $M \in U(T)$.

The proof we use is a more refined version of the proof of Theorem 6.28. Thus we produce $M$ from $A$ in stages. We first use a 1-step construction to embed $A$ into some $A' \in S(T)$ which is only partly existentially universal (in an appropriate sense). We then iterate this construction to accumulates this partial saturation into a full saturation of the required kind. These stages are analogous to Lemma 6.27 and Theorem 6.28.

We need to gain some control over the size of the constructed $M$. In particular, we wish to keep $M$ as small as possible. With the construction of existentially closed structures, this is not a problem. Once we have one such structure, we may take an elementary substructure to produce a smaller one. With existentially universal structures this is not possible (since we need a preservation property that is stronger than the one provided by an elementary embedding). Thus, we must take more care with the construction.

In the 1-step construction $A \hookrightarrow A'$ we need to calculate the size of $A'$ in comparison with the size of $A$. Let's see how we do that.

We work relative to a theory $T$ which we assume is formalized in some language $L$. The cardinality $\lambda = |L|$ of this language is an important parameter in the construction. This cardinal $\lambda$ has an impact on the size of the constructed structure. In this first instance it will do know harm if you assume that $\lambda = \aleph_0$. Also, for many applications we do need a countable language.

We assume given some $A \in S(T)$, and we want to produce some $A \subseteq A' \in S(T)$ which realizes many $\exists_1$-types. Each such type has the form

$$\Theta(a, v)$$

where $\Theta(u, v)$ is a pure $\exists_1$-type, and the point $a$ come from $A$ (not yet the whole of $A'$).

How many such types are there?

Since the $|L| = \lambda$, there are no more than $2^\lambda$ pure types $\Theta(u, v)$. Each one of these may give many different types $\Theta(a, v)$ as $a$ varies through $A$. Suppose

$$\lambda \leq \kappa \quad |A| \leq 2^\kappa$$

for some infinite cardinal $\kappa$. Then there are no more than

$$2^\lambda \cdot 2^\kappa = 2^{\lambda + \kappa} = 2^\kappa$$

types $\Theta(a, v)$. Using this we will arrange that $|A'| \leq 2^\kappa$. In particular, as we accumulate

$$A \hookrightarrow A' \hookrightarrow A'' \hookrightarrow A''' \hookrightarrow \cdots$$

at each step the structure produced will have size no more that $2^\kappa$, and hence the final structure will have size no more that $2^\kappa \cdot \aleph_0 = 2^\kappa$. 
7.18 LEMMA. (The 1-step construction) Let $T$ be a theory in a language of cardinality $\lambda$. Let $\mathfrak{A} \in S(T)$ with $|\mathfrak{A}| \leq 2^\kappa$ for some cardinal $\kappa$ with $\lambda \leq \kappa$. Then there is a structure $\mathfrak{A}' \in S(T)$ with

$$\mathfrak{A} \subseteq \mathfrak{A}' \quad |\mathfrak{A}'| \leq 2^\kappa$$

and such that for each $\exists_1$-type $\Theta(a, v)$ with parameters from $\mathfrak{A}$, if this type is $T$-consistent over $\mathfrak{A}'$, then it is already realized in $\mathfrak{A}'$.

**Proof.** Consider all possible $\exists_1$-types with parameters from $\mathfrak{A}$. By the discussion above, we know there are no more than $2^\kappa$ such types. Let

$$\Theta = \{\Theta_i \mid i < 2^\kappa\}$$

by an ordinal enumeration of these types. It doesn’t matter if there are some repetitions in this enumeration. Also, each $\Theta_i$ will contain certain parameters and free variables, and these may change with $i$, but we do not need the details of this dependence.

We produce $\mathfrak{A}'$ as the union of a long ascending chain

$$\mathcal{A} = \{\mathfrak{A}_i \mid i < 2^\kappa\}$$

of submodels of $T$. We generate this chain by recursion along the ordinals $i$ up to $2^\kappa$. For each $i < 2^\kappa$, the recursion step $\mathfrak{A}_i \mathrel{\longleftarrow} \mathfrak{A}_{i+1}$ deals with the type $\Theta_i$.

Here is what we arrange.

1. $\mathfrak{A}_0 = \mathfrak{A}$.
2. $i < j < 2^\kappa \implies \mathfrak{A}_i \subseteq \mathfrak{A}_j$.
3. $|\mathfrak{A}_i| \leq 2^\kappa$ for each $i < 2^\kappa$.
4. If $l < 2^\kappa$ is a limit ordinal, then $\mathfrak{A}_l = \bigcup\{\mathfrak{A}_i \mid i < l\}$.
5. For each $i < 2^\kappa$, if the type $\Theta_i$ is $T$-consistent over $\mathfrak{A}_i$, then it is realized in $\mathfrak{A}_{i+1}$.

Before we generate this chain $\mathcal{A}$, let’s see why

$$\mathfrak{A}' = \bigcup \mathcal{A}$$

has the required properties.

Firstly, this $\mathfrak{A}'$ is the union of an ascending chain of submodels of $T$, and hence itself is a submodel of $T$.

Secondly, by (2) and since the length of the chain is $2^\kappa$, we have $|\mathfrak{A}'| \leq 2^\kappa$.

Thirdly, it has the required partial saturation property. Consider any $\exists_1$-type $\Theta(a, v)$ with parameters from $\mathfrak{A}$. Suppose also that this type is $T$-consistent over $\mathfrak{A}'$. By the choice of $\Theta$, we have $\Theta = \Theta_i$ for some $i < 2^\kappa$, and then $\Theta_i$ is $T$-consistent over $\mathfrak{A}_i$. By clause (4) this $\Theta_i$ is realized in $\mathfrak{A}_{i+1}$ and hence, since $\Theta = \Theta_i$ is a $\exists_1$-type, it is realized in $\mathfrak{A}'$.

It remains to generate the chain $\mathcal{A}$. We do this by recursion along the ordinals up to $2^\kappa$.

For the base case, $i = 0$, we set $\mathfrak{A}_0 = \mathfrak{A}$, to obtain (0). Since we are given $|\mathfrak{A}| \leq 2^\kappa$, we have (2) (for this $i$).
For the recursion step, \( i \mapsto i + 1 \), we look at the type \( \Theta_i \). There are two sub-cases.

If \( \Theta_i \) is not \( T \)-consistent over \( \mathfrak{A}_i \), then we set \( \mathfrak{A}_{i+1} = \mathfrak{A}_i \). In this sub-case, properties (1), (2), and (4) are immediate.

Suppose \( \Theta_i \) is \( T \)-consistent over \( \mathfrak{A}_i \). There is some \( \mathfrak{A}_i \subseteq \mathfrak{B} \in \mathcal{S}(T) \) in which \( \Theta_i \) is realized. By taking a suitable elementary substructure, we may suppose \( |\mathfrak{B}| = |\mathfrak{A}_i| \leq 2^\kappa \).

We let \( \mathfrak{A}_{i+1} \) be such a \( \mathfrak{B} \), and so preserve properties (1), (2), and (4).

For the leap to a limit ordinal \( l < 2^\kappa \), we must set \( \mathfrak{A}_l = \bigcup\{\mathfrak{A}_i \mid i < l\} \) by (3). Property (1) is preserved, and (4) is vacuous, so we must check (3). But, using (2) we have

\[
|\mathfrak{A}_l| \leq 2^\kappa \cdot |l| = 2^\kappa
\]

(since \(|l| < 2^\kappa \)), as required.

This completes the construction, and the whole proof.

Before we continue it is worth looking at the restrictions on the size of the constructed \( \mathfrak{A}' \). Let’s look at the simplest case where \( \lambda = \aleph_0 \) and the given structure \( \mathfrak{A} \) is countable, so that \(|\mathfrak{A}| \leq \aleph_0 < 2^{\aleph_0} \). Perhaps this strict comparison can help us to produce a smaller extension \( \mathfrak{A}' \in \mathcal{S}(T) \). However, the problem is not the size of \( \mathfrak{A} \) but the potential number of types that have to be realized. This is still \( 2^{\aleph_0} \), so the size of the first step structure \( \mathfrak{A}' \) still could be \( 2^{\aleph_0} \). At the next step we move from \( \mathfrak{A}' \) to \( \mathfrak{A}'' \), so we are back in the situation of Lemma 7.18 with \( \kappa = \aleph_0 \). We will return to this discussion later.

7.19 THEOREM. (The accumulation construction) Let \( T \) be a theory in a language of cardinality \( \lambda \). Let \( \mathfrak{A} \in \mathcal{S}(T) \) with \(|\mathfrak{A}| \leq 2^\kappa \) for some cardinal \( \kappa \) with \( \lambda \leq \kappa \). Then there is a structure \( \mathfrak{M} \in \mathcal{U}(T) \) with \( \mathfrak{A} \subseteq \mathfrak{M} \) and \(|\mathfrak{M}| \leq 2^\kappa \).

Proof. Given a structure \( \mathfrak{A} \in \mathcal{S}(T) \) we may iterate the 1-step construction of Lemma 7.18 to produce an \( \omega \)-chain

\[
\mathfrak{A} = \mathfrak{A}_0 \subseteq \mathfrak{A}_1 \subseteq \cdots \subseteq \mathfrak{A}_i \subseteq \cdots \quad (i < \omega)
\]

of submodels of \( T \) where, for each step \( i \), the structure \( \mathfrak{A}_{i+1} \) realizes certain types taken from \( \mathfrak{A}_i \). Let

\[
\mathfrak{M} = \bigcup\{\mathfrak{A}_i \mid i < \omega\}
\]

so that \( \mathfrak{A} \subseteq \mathfrak{M} \in \mathcal{S}(T) \). At each step we have

\[
|\mathfrak{A}_i| \leq 2^\kappa
\]

so that \(|\mathfrak{M}| \leq 2^\kappa \). Thus it suffices to show that \( \mathfrak{M} \in \mathcal{U}(T) \).

Consider any \( \exists_1 \)-type \( \Theta(a, v) \) with parameters \( a \) from \( \mathfrak{M} \) and which is \( T \)-consistent over \( \mathfrak{M} \). This point \( a \) confines just finitely many elements of \( \mathfrak{M} \), and hence \( a \) is a point of \( \mathfrak{A}_i \) for some \( i < \omega \). Furthermore, the type \( \Theta \) is \( T \)-consistent over \( \mathfrak{A}_i \). By construction, the \( \exists_1 \)-type \( \Theta \) is realized in \( \mathfrak{A}_{i+1} \), and hence is realized in \( \mathfrak{M} \), as required.

Suppose \( T \) is a theory in a countable language, and suppose \( \mathfrak{A} \in \mathcal{S}(T) \) is countable. Theorem 7.19 gives us some \( \mathfrak{A} \subseteq \mathfrak{M} \in \mathcal{U}(T) \) with \(|\mathfrak{M}| \leq 2^{\aleph_0} \). In terms of size, is this the best we can do? The following example shows that, in general, it is.
7.3. Existence of existentially universal structures

7.20 EXAMPLE. Consider the binary splitting tree as a structure

\[ \mathcal{T} = (\Psi, \leq, \bot, S_0, S_1) \]

where \( \Psi \) is the set of all node; \( \leq \) is the comparison of these nodes; and \( \bot \) is the empty node, the bottom of the poset. The two 1-placed operation symbols \( S_0 \) and \( S_1 \) are the left and right successor operations. Thus

\[ S_0 a = a0 \quad S_1 a = a1 \]

for each node \( a \). Notice that each node \( a \) has a canonical name

\[ a = S_{i(0)}(\cdots (S_{i(0)}) \cdots) \]

where each \( S_{i(\cdot)} \) is \( S_0 \) or \( S_1 \), as appropriate. We often think of this as a list

\[ i(0)i(1) \cdots i(n) \]

of zeros and ones.

Let \( T = Th(\mathcal{T}) \). This is the **theory of two successor functions**. We need not write down the axioms for this theory, but we should note that the sentence

\[ (\forall u, v, w)(u \leq w \land v \leq w \rightarrow (u \leq v) \lor (v \leq u)) \]

is in \( T \).

The structure \( \mathcal{T} \) is countable, and \( \mathcal{T} \subseteq \mathcal{M} \) for some \( \mathcal{M} \in \mathcal{U}(T) \). How big must \( \mathcal{M} \) be? Consider any branch \( p \) of \( \Psi \). Let

\[ \Theta_p(w) = \{(a \leq w) | a < p\} \]

to produce a quantifier-free pure type in the single variable \( w \). Each \( a \) here is a node of \( \Psi \) and so is a term of the underlying language.

This type is finitely satisfiable in \( \mathcal{T} \), and hence in \( \mathcal{M} \). Thus it is realized in some elementary extension of \( \mathcal{M} \). But \( \mathcal{M} \in \mathcal{U}(T) \) and hence the type is realized in \( \mathcal{M} \).

For each branch \( p \) of \( \Psi \) the type \( \Theta_p \) is realized by some element of \( \mathcal{M} \). We check that no two branches are realized by the same element, and hence \( |\mathcal{M}| \geq 2^{\aleph_0} \).

Consider distinct branches \( p, q \) of \( \Psi \), and suppose

\[ \mathcal{M} \models \Theta_p(m) \quad \mathcal{M} \models \Theta_q(m) \]

for some \( m \in \mathcal{M} \). Since \( p \neq q \), there are nodes \( a, b \in \Psi \) such that

\[ a < p \quad a \not< q \quad b < q \quad b \not< p \]

hold. But

\[ a \leq m \quad b \leq m \]

in \( \mathcal{M} \), so that

\[ a \leq b \quad \text{or} \quad b \leq a \]

and hence

\[ a \leq q \quad \text{or} \quad b \leq p \]

neither of which can hold. \( \blacksquare \)
The problem with the size of an existentially universal structure is the number of types it must realize. If we can keep this small, then we have a chance of keeping down the size of any existentially universal structure we may construct.

Consider the case of a countable language, and have another look at the calculations just before Lemma 7.18. Suppose we want to embed some countable $A \in S(T)$ into some $M \in U(T)$. The number of types to be considered could be $2^{\aleph_0}$, and then Example 7.20 show that $|M| = 2^{\aleph_0}$ is the best we can achieve. However, sometimes we know there is a smaller number of types, and then we can do better.

Consider how we produce $A'$ from $A$, as in the proof of Lemma 7.18. We have to realize in $A'$ each $\exists_1$-type $\Theta(a, v)$ which is $T$-consistent over $A'$ but has parameters from $A$. Each such type is realized in some existentially closed extension of $A'$ and so is a subtype of some $\exists_1$-type which is maximal over $T$. Thus, in the construction, it suffices to consider only types $\Theta$ which arise by instantiation from maximal types. This in itself doesn’t bring down the number of types, but we know a method that does.

Recall that for each list $w$ of variable

$$\exists_1^{\text{max}}(T, w)$$

is the set of $\exists_1$-types in the batch $w$ each of which is maximally consistent over $T$.

7.21 LEMMA. (The refined 1-step construction) Let $T$ be a theory in a countable language, and suppose $\exists_1^{\text{max}}(T, w) \leq \aleph_0$ for each batch $w$ of variable. Then for each countable $A \in S(T)$ there is a countable $A' \in S(T)$ with $A \subseteq A'$ and such that for each $\exists_1$-type $\Theta(a, v)$ with parameters from $A$, if this type is $T$-consistent over $A'$, then it is already realized in $A'$.

Proof. We repeat the construction of the proof of Lemma 7.21. But now we have a countable enumeration

$$\Theta = \{\Theta_i | i < \omega\}$$

of the types that have to be handle. Thus we can produce $A'$ as the union of an ascending $\omega$-chain

$$A = \{A_i | i < \omega\}$$

of structures each of which is countable. Hence $A'$ is countable.

The accumulation construction is always an $\omega$-iteration of the 1-step construction. Thus an iteration of Lemma 7.21 gives the following.

7.22 THEOREM. (The refined accumulation construction) Let $T$ be a theory in a countable language, and suppose $\exists_1^{\text{max}}(T, w) \leq \aleph_0$ for each list $w$ of variable. Then for each countable $A \in S(T)$ there is a countable $M \in U(T)$ with $A \subseteq M$.

What is the use of this? In other words, how can we keep down the size of $\exists_1^{\text{max}}(T, w)$? That is one of the topics of the next chapter.

Exercises

Some needed
7.4 Atomicity

So far in this chapter we have been looking at structures that are rather large, in a certain sense. In this section we go to the other end of the scale. We look at structures that are rather small in the sense that each such structure contains a point only when it has to. The type of each point is controlled by a single formula, and so it is easy to say that such a point must exist (by existentially quantifying out the variables of this controlling formula).

We will consider two extreme versions, the $\omega$-version and the 0-version. The $\omega$-version is concerned with elementary embeddings and the like, whereas the 0-version is concerned with embeddings and the like.

For the $\omega$-version we analyse a complete theory $T$ and can work entirely within the class $Md(T)$ of models of $T$.

For the 0-version we analyse a theory $T$ which, in practice, should have JEP. We use associated gadgets such as $T^0$ and $E(T)$, and as usual we work within the class $S(T)$ of submodels of $T$.

It is the $\omega$-version that is most often described in an account of model theory. The 0-version is slightly more complicated and is not often described. Furthermore, once we have the 0-version the $\omega$-version is easy to obtain (but not the other way round). Thus here we concentrate mainly on the 0-version.

There is also an $n$-version for each $n \in \mathbb{N}$. But that material is a more or less routine generalization of the 0-version, and adds little to our understanding of the notions involved.

7.23 DEFINITION. Let $T$ be a theory.

($\omega$) A formula $\theta$ is complete over $T$ if (it is consistent with $T$) and exactly one of

$$T \vdash \theta \rightarrow \psi \quad T \vdash \theta \rightarrow \neg \psi$$

holds for each formula $\psi$ (with $\partial \psi \subseteq \partial \theta$).

(0) An $\exists_1$-formula $\theta$ is $\exists_1$-complete over $T$ if it is consistent with $T$ and

$$T \cup \{\theta, \psi_1\} \text{ is consistent } \quad T \cup \{\theta, \psi_2\} \text{ is consistent}$$

holds for all $\exists_1$-formulas $\psi_1, \psi_2$ (with $\partial \psi_1 \cup \partial \psi_2 \subseteq \partial \theta$).  

Notice that if a formula $\theta$ is complete over a theory $T$ then it must be consistent with $T$, for otherwise we have

$$T \vdash \theta \rightarrow \psi$$

for every formula $\psi$.

The $\omega$-version of this notion can be rephrased to look more like the 0-version. However, the phrasing above is the standard way to describe this notion. The 0-version can not be rephrased to look like the $\omega$-version. This is because the negation of a $\exists_1$-formula need not be a $\exists_1$-formula.

There are various characterizations of $\exists_1$-completeness, and we need some of these. We use the 0-companion $T^0$ of the parent theory $T$. Recall that an $\exists_2$-formula $\chi(v)$ is consistent with $T^0$ exactly when $\mathfrak{A} \models \chi(a)$ for some $\mathfrak{A} \in \mathcal{E}(T)$ and some point $a$ of $\mathfrak{A}$.

\footnote{Exercise: Prove this.}

*1 This should have been an earlier exercise.*
7.24 LEMMA. Let $T$ be a theory, and let $\theta$ be an $\exists_1$-formula consistent with $T$. The following are equivalent.

(i) The formula $\theta$ is $\exists_1$-complete over $T$.

(ii) We have

$$T^0 \cup \{\theta, \phi\} \text{ is consistent } \implies T \vdash \theta \rightarrow \phi$$

for each $\forall_1$-formula $\phi$ (with $\partial \phi \subseteq \partial \theta$).

(iii) We have

$$T \cup \{\theta, \psi\} \text{ is consistent } \implies T^0 \vdash \theta \rightarrow \psi$$

for each $\exists_1$-formula $\psi$ (with $\partial \psi \subseteq \partial \theta$).

Proof. (i) $\Rightarrow$ (ii). Assuming (i), consider any $\forall_1$-formula $\phi$ (with $\partial \phi \subseteq \partial \theta$) where $T^0 \cup \{\theta, \phi\}$ is consistent. Since $\theta \land \phi$ is a $\exists_2$-formula, we have

$$A \models \theta(a) \quad A \models \phi(a)$$

for some $A \in \mathcal{E}(T)$ and some point $a$ from $A$. By Lemma 6.26 the second of these gives some $\exists_1$-formula $\psi$ such that

$$A \models \psi(a) \quad T \vdash \psi \rightarrow \phi$$

hold. In particular,

$$T \cup \{\theta, \psi\}$$

is consistent, but

$$T \cup \{\psi, \neg \phi\}$$

is not consistent. Thus, since $\neg \phi$ is a $\exists_1$-formula, the assumption (i) shows that

$$T \cup \{\theta, \neg \phi\}$$

is not consistent, and hence

$$T \vdash \theta \rightarrow \phi$$

as required.

(ii) $\Rightarrow$ (iii). Assuming (ii) consider any $\exists_1$-formula $\psi$ (with $\partial \psi \subseteq \partial \theta$) where $T \cup \{\theta, \psi\}$ is consistent. If $T^0 \not\vdash \theta \rightarrow \psi$, then $T^0 \cup \{\theta, \neg \psi\}$ is consistent, and so (ii) gives

$$T \vdash \theta \rightarrow \neg \psi$$

(since $\neg \psi \in \forall_1$). This is not so, and hence we have $T^0 \vdash \theta \rightarrow \psi$, as required.

(iii) $\Rightarrow$ (i). Since $\theta$ is consistent with $T$ we have $A \models \theta(a)$ for some $A \models T$ and some point $a$ from $A$. Consider

$$A \subseteq B \subseteq C$$

where $B \models T^0$ and $C \models T$. Note that $B \models \theta(a)$ and $C \models \theta(a)$. Let $\Psi$ be the $\exists_1$-type of $a$ in $C$. For each $\exists_1$-formula $\psi$, the assumption (iii) gives

$$T \cup \{\theta, \psi\} \text{ consistent } \implies T^0 \vdash \theta \rightarrow \psi \implies B \models \psi(a) \implies C \models \psi(a) \implies \psi \in \Psi$$

which leads to (i).

Almost invariably we use the notion of an $\exists_1$-complete formula only for a theory with $JEP$. We use the notion of a complete formula only for a complete theory.
7.25 DEFINITION. Let \( T \) be a theory.

(\( \omega \)) A structure \( \mathfrak{A} \) is atomic for \( T \) if \( \mathfrak{A} \models T \) and for each point \( a \) of \( \mathfrak{A} \) there is a formula \( \theta \) which is complete over \( T \) with \( \mathfrak{A} \models \theta(a) \).

(0) A structure \( \mathfrak{A} \) is \( \exists_1 \)-atomic for \( T \) if \( \mathfrak{A} \models T \) and for each point \( a \) of \( \mathfrak{A} \) there is a formula \( \theta \) which is \( \exists_1 \)-complete over \( T \) with \( \mathfrak{A} \models \theta(a) \).

Let \( \mathcal{A}(T) \) be the class of structures which are \( \exists_1 \)-atomic for \( T \).

Notice again how the companion \( T \) is used in the 0-version. This ensures that an \( \exists_1 \)-atomic structure is existentially closed. In fact, as we will see later, it is rather a special kind of existentially closed structure.

Unlike most other classes we associate with a theory \( T \), the class \( \mathcal{A}(T) \) can be empty. At the end of this section we will see how to arrange that \( \mathcal{A}(T) \) is non-empty.

7.26 LEMMA. Let \( T \) be a theory. For each structure \( \mathfrak{A} \) the following are equivalent.

(i) \( \mathfrak{A} \) is \( \exists_1 \)-atomic for \( T \).

(ii) \( \mathfrak{A} \in \mathcal{E}(T) \) and for each point \( a \) of \( \mathfrak{A} \) there is an \( \exists_1 \)-formula \( \theta \) such that

\[
\mathfrak{A} \models \theta(a) \quad T^0 \vdash \theta \rightarrow \bigwedge \Sigma
\]

where \( \Sigma \) is the \( \exists_1 \)-type of \( a \) in \( \mathfrak{A} \).

(iii) \( \mathfrak{A} \in \mathcal{S}(T) \) and for each point \( a \) of \( \mathfrak{A} \) there is an \( \exists_1 \)-formula \( \theta \) such that

\[
\mathfrak{A} \models \theta(a) \quad T \vdash \theta \rightarrow \bigwedge \Pi
\]

where \( \Pi \) is the \( \forall_1 \)-type of \( a \) in \( \mathfrak{A} \).

Proof. (i)\( \Rightarrow \) (ii). Assuming (i) consider any point \( a \) of \( \mathfrak{A} \), and let \( \theta \) be any formula which is \( \exists_1 \)-complete over \( T \) and \( \mathfrak{A} \models \theta(a) \). Let \( \Sigma \) and \( \Pi \) be, respectively, the \( \exists_1 \)-type of \( a \) and the \( \forall_1 \)-type of \( a \) in \( \mathfrak{A} \). Since \( \mathfrak{A} \models T^0 \) we see that

\[
T^0 \cup \{ \theta \} \cup \Sigma \cup \Pi
\]

is consistent. Thus we have

\[
T \vdash \theta \rightarrow \bigwedge \Pi \quad T^0 \vdash \theta \rightarrow \bigwedge \Sigma
\]

by Lemma 7.24. The first of these shows that \( \mathfrak{A} \in \mathcal{E}(T) \) and the second completes the proof of (ii).

(ii)\( \Rightarrow \) (iii). Assuming (ii) consider any point \( a \) of \( \mathfrak{A} \), and let \( \theta \) be the \( \exists_1 \)-formula given by (ii). Let \( \Sigma \) and \( \Pi \) be, respectively, the \( \exists_1 \)-type of \( a \) and the \( \forall_1 \)-type of \( a \) in \( \mathfrak{A} \). We have

\[
T \vdash \bigwedge \Sigma \rightarrow \bigwedge \Pi
\]

since \( \mathfrak{A} \in \mathcal{E}(T) \). By (ii) we have

\[
T^0 \vdash \theta \rightarrow \bigwedge \Sigma
\]

and hence

\[
T \vdash \theta \rightarrow \bigwedge \Pi
\]
to complete the proof of (iii).

(iii)⇒(i). Assuming (iii) observe first of all that $\mathfrak{A} \in \mathcal{E}(T)$ and hence $\mathfrak{A} \models T^0$. Consider any point $a$ of $\mathfrak{A}$ and let $\theta$ be the $\exists_1$-formula given by (iii). We show that $\theta$ is $\exists_1$-complete over $T$. To do this we verify the property of Lemma 7.24(ii).

Consider any $\forall_1$-formula $\phi$ (with $\partial \phi \subseteq \partial \theta$ and) where $T^0 \cup \{ \theta, \phi \}$ is consistent. There is some $\mathfrak{B} \in \mathcal{E}(T)$ and some point $b$ of $\mathfrak{B}$ with $\mathfrak{B} \models \theta(b) \land \phi(b)$. The assumed property (iii) ensures that $(\mathfrak{A}, a) \equiv (\forall_1) (\mathfrak{B}, b)$ and hence

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

(since $\mathfrak{B} \in \mathcal{E}(T)$). Thus we have $\mathfrak{A} \models \phi(a)$ (so that $\phi \in \Pi$), and hence $T \vdash \theta \rightarrow \phi$ by a second use of (iii).

It is perhaps a little surprising but atomic structures have strong back-and-forth properties. The following result and its corollary should be compared with Theorem 7.8 and its corollary.

7.27 THEOREM. Let $T$ be a theory, and let $\mathfrak{A}, \mathfrak{B} \in \mathcal{A}(T)$. The set of pairs $(a, b)$ of points $a$ from $\mathfrak{A}$ and $b$ from $\mathfrak{B}$ for which

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

is a back-and-forth system for $\mathfrak{A}, \mathfrak{B}$.

Proof. Consider any pair $(a, b)$ with

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

and any element $x$ of $\mathfrak{A}$. Let $\Sigma(u, v)$ be type of the extended point $a^\chi$ of $\mathfrak{A}$. Thus $u$ is a list of variable matching $a$, and $v$ is a single variable. By Lemma 7.26(ii) there is an $\exists_1$-formula $\theta(u, v)$ such that

$\mathfrak{A} \models \theta(a, x) \quad T^0 \vdash \theta \rightarrow \bigwedge \Sigma$

holds. In particular, we have $\mathfrak{A} \models (\exists v) \theta(a, v)$ and hence $\mathfrak{B} \models (\exists v) \theta(b, v)$ (by the relationship between $a$ and $b$). This provides an element $y$ of $\mathfrak{B}$ with $\mathfrak{B} \models \theta(b, y)$. Since $\mathfrak{B} \models T^0$, this gives $\mathfrak{B} \models \Sigma(b, y)$, so that

$$(\mathfrak{A}, a, x) \equiv (\exists_1) (\mathfrak{B}, b, y)$$

and hence

$$(\mathfrak{A}, a, x) \equiv_1 (\mathfrak{B}, b, y)$$

(since $\mathfrak{A}, \mathfrak{B} \in \mathcal{E}(T)$).

This, with a similar argument for the other direction, verifies the back-and-forth property.

This has an immediate consequence for the family of $\exists_1$-atomic structures.
7.28 COROLLARY. Let $T$ be a theory. We have

$$A \equiv_1 B \implies A \equiv_p B$$

for each $A, B \in A(T)$.

We know that the two relations $\equiv_p$ and $\equiv_\infty$ agree on countable structures, and hence we have the following uniqueness property.

7.29 COROLLARY. Let $T$ be a theory with JEP. Then, up to isomorphisms, there is at most one countable structure which is $\exists_1$-atomic for $T$.

We now come to the crucial result, the existence of atomic structures. In general for a theory an atomic structure (of either kind) need not exist. We need to impose a syntactic condition on the theory.

7.30 DEFINITION. $(\omega)$ A theory $T$ is atomic if it is complete and for each formula $\psi$ which is consistent with $T$ we have

$$T \vdash \theta \to \psi$$

for some formula $\theta$ which is complete over $T$.

$(0)$ A theory $T$ is $\exists_1$-atomic if it has JEP and for each $\exists_1$-formula $\psi$ which is consistent with $T$ we have

$$T^0 \vdash \theta \to \psi$$

for some formula $\theta$ which is $\exists_1$-complete over $T$. ■

One connection between atomic structure and atomic theories is easy to obtain.

7.31 LEMMA. Let $T$ be a theory with JEP and suppose there is a structure which is $\exists_1$-atomic for $T$. Then $T$ is $\exists_1$-atomic.

Proof. Suppose the structure $\mathfrak{A}$ is $\exists_1$-atomic for $T$, and consider any $\exists_1$-formula $\psi$ which is consistent with $T$. Since $T$ has JEP we know that $\psi$ is realized in $\mathfrak{A}$, so there is some point $a$ of $\mathfrak{A}$ with $\mathfrak{A} \models \psi(a)$. By Lemma 7.26(ii) there is some formula $\theta$ which is $\exists_1$-complete over $T$ with $\mathfrak{A} \models \theta(a)$ and $T^0 \vdash \theta \to \psi$. This show that $T$ is $\exists_1$-atomic. ■

Notice that this result does not require any countability neither of the language nor the structure. Of course, we want to go the other way and that does require a countability restriction.

We wish to improve Lemma 7.31 to an equivalence, but to that we need to assume that the underlying language is countable.

So far this chapter has been an introduction to the saturation technique. To improve Lemma 7.31 we use the omitting types technique, and this last party is designed as an introduction to that technique.

7.32 DEFINITION. Let $T$ be a theory. For each batch $v$ of formulas let $\Omega(T)(v)$ be the set of all formulas

$$\neg \theta$$

where $\theta$ is $\exists_1$-complete over $T$ with $\partial \theta = v$. Thus $\Omega(T)(v)$ is an $\forall_1$-type.

Let $\Omega(T)$ be the set of all such $\forall_1$-types $\Omega(T)(v)$ (for varying batches $v$). ■
As we will see from the proof, the following characterization is little more than a rephrasing of Definition 7.25.

7.33 LEMMA. Let $T$ be a theory. For each structure $\mathfrak{A}$ the following are equivalent.

(i) $\mathfrak{A}$ is $\exists_1$-atomic for $T$.

(ii) $\mathfrak{A} \in E(T)$ and $\mathfrak{A}$ omits each type in $\Omega(T)$.

Proof. (i)$\Rightarrow$(ii). Suppose $\mathfrak{A}$ is $\exists_1$-atomic for $T$. Then certainly $\mathfrak{A} \in E(T)$. Consider any batch $v$ of variables. We must show that $\mathfrak{A}$ omits $\Omega(T)(v)$.

To this end consider any point $a$ of $\mathfrak{A}$ which matches $v$. By Definition 7.25 there is some $\exists_1$-formula $\theta(v)$ which is $\exists_1$-complete over $T$ and $\mathfrak{A} \models \theta(a)$. We then have $\neg \theta \in \Omega(T)(v)$, and hence $\mathfrak{A}$ does not realize $\Omega(T)(v)$ at $a$. Since this is true for each point $a$ of $\mathfrak{A}$, we see that $\mathfrak{A}$ omits $\Omega(T)(v)$.

(ii)$\Rightarrow$(i). Assuming (ii) we certainly have $\mathfrak{A} \models T^0$. Consider any point $a$ of $\mathfrak{A}$. Let $v$ be a batch of variable matching $a$. Since $\mathfrak{A}$ omits $\Omega(T)(v)$ there is some formula $\theta(v)$ with $\neg \theta \in \Omega(T)(v)$ and with $\mathfrak{A} \models \theta(a)$. Since this $\theta$ is $\exists_1$-complete over $T$, we have verified the requirements of By Definition 7.25. ■

Now we are almost ready for the crunch. We need just one more bit of terminology. You should compare the following notion with that in part (iii) of Lemma 7.26.

7.34 DEFINITION. We say a $\forall_1$-type $\Pi$ is $\exists_1$-principal over a theory $T$ if we have

$$T \vdash \theta \rightarrow \bigwedge \Pi$$

for some $\exists_1$-formula $\theta$ with $\partial \theta = \partial \Pi$ and which is consistent over $T$. ■

To obtain a converse of Lemma 7.31 we invoke an omitting types theorem. More precisely we apply the following.

7.35 THEOREM. Let $T$ be a theory in a countable language. Let $\Pi$ be a countable collection of $\forall_1$-types each of which is not $\exists_1$-principal over $T$. Then there is a countable structure $\mathfrak{A} \in E(T)$ which omits $\Pi$.

We can not prove this here, that must wait until [**section on omitting types**]. The idea is that the last part of this section is an introduction to the omitting types technique.

7.36 THEOREM. Let $T$ be a theory in a countable language, and suppose $T$ has JEP. The following are equivalent.

(i) There is a countable structure which is $\exists_1$-atomic for $T$.

(ii) The theory $T$ is $\exists_1$-atomic.
7.4. Atomicity

Proof. (i)⇒(ii). This is given by Lemma 7.31

(ii)⇒(i). Suppose \( T \) is \( \exists_1 \)-atomic. For each finite batch \( v \) of variable we have an \( \forall_1 \)-type \( \Omega(T)(v) \), as in Definition 7.32. Since the language is countable and there are only countably many batches of variables, the whole family \( \Omega(T) \) of these types is countable. We apply Theorem 7.35 to this family \( \Omega(T) \) and so obtain some countable \( A \in \mathcal{E}(T) \) which omits each member of \( \Omega(T) \). By Lemma 7.33 such a structure is \( \exists_1 \)-atomic for \( T \).

Consider any member \( \Omega(T)(v) \) of \( \Omega(T) \). We must show that \( \Omega(T)(v) \) is not \( \exists_1 \)-principal over \( T \).

By way of contradiction, suppose \( \Omega(T)(v) \) is \( \exists_1 \)-principal over \( T \). Thus we have
\[ T \vdash \psi \to \bigwedge \Omega(T)(v) \]
for some some \( \exists_1 \)-formula \( \psi(v) \) which is consistent with \( T \). This gives
\[ T^0 \vdash \psi \to \bigwedge \Omega(T)(v) \]
since \( T \) and \( T^0 \) are companions.

By (ii) we have
\[ T^0 \vdash \theta \to \psi \]
for some formula \( \theta(v) \) which is \( \exists_1 \)-complete over \( T \).

But now we have
\[ T^0 \vdash \theta \to \bigwedge \Omega(T)(v) \]
and hence
\[ T \vdash \theta \to \bigwedge \Omega(T)(v) \]
since \( T \) and \( T^0 \) are still companions.

By the construction of \( \Omega(T)(v) \) we have \( \neg \in \theta \Omega(T)(v) \), so that
\[ T \vdash \theta \to \neg \theta \]
to give \( T \vdash \neg \theta \), which is a contradicts since \( T \cup \{ \theta \} \) is consistent. \( \blacksquare \)

This is the main existence result for the 0-version. The corresponding existence result for the \( \omega \)-version is similar.

7.37 THEOREM. Let \( T \) be a complete theory in a countable language. The following are equivalent.

(i) There is a countable structure which is atomic for \( T \).

(ii) The theory \( T \) is atomic.

The proof of this follows the same pattern as the proof of Theorem 7.36. The difference is that we no longer have any restrictions on the quantifier complexity of formulas and types. To help with this, all embeddings are elementary, and so formulas are preserved when we need them to be. At the heart of the proof is an application of the omitting types result for full types.

To conclude this section and chapter we end with a couple of questions.

Let \( T \) be a theory which is \( \exists_1 \)-atomic (in a countable language). In particular, \( T \) has JEP. By combining this Corollary with Theorem 7.36 we see that \( T \) has a unique countable \( \exists_1 \)-atomic submodel. This must hold some privileged position in the spectrum of all submodels. What is this? Let \( T^a = \text{Th}(\mathcal{A}(T)) \). By Corollary 7.28 we see that \( T^a \) is complete, and \( T^c \subseteq T^a \) (since \( \mathcal{A}(T) \subseteq \mathcal{E}(T) \)). What is this theory \( T^a \)?
7. Small and large structures

Exercises

7.6 Let $T$ be an arbitrary theory.
    Show that a formula $\theta$ is complete over $T$ precisely when it is consistent with $T$ and we have
    
    $T \cup \{\theta, \psi_1\}$ is consistent
    $T \cup \{\theta, \psi_2\}$ is consistent
    \[\implies\]
    $T \cup \{\psi_1, \psi_2\}$ is consistent
    
    for all formulas $\psi_1, \psi_2$ (with $\partial \psi_1 \cup \partial \psi_2 \subseteq \partial \theta$).

7.7 Let $\theta$ be an $\exists_1$-formula consistent with a theory $T$. Show that $\theta$ is $\exists_1$-complete over $T$ if and only if
    
    $T \models \phi_1 \lor \phi_2 \implies T \models \theta \rightarrow \phi_1$ or $T \models \theta \rightarrow \phi_2$
    
    holds for all $\forall_1$-formulas $\phi_1, \phi_2$.

7.8 Show that if $T$ is $\exists_1$-atomic, then for each $\forall_1$-formula $\phi(v)$ which is consistent with $T$, there is a formula $\theta(v)$ which is $\exists_1$-complete over $T$ for which $T \models \theta \rightarrow \phi$.

7.9 (a) For an arbitrary $n < \omega$ write down the notion of a formula being $\exists_{n+1}$-complete over a theory $T$.
    (b) Show a formula $\theta$ is complete over a theory $T$ if and only if it is $\exists_{n+1}$-complete over $T$ for all sufficiently large $n < \omega$.

7.10 Prove Theorem 7.37.

7.11 In Corollary 7.29, why is it necessary to assume that the theory has JEP?

7.12 Let $T$ be a theory in a countable language, and suppose $T$ has JEP. Let $\mathfrak{A}$ be the unique countable $\exists_1$-atomic structure for $T$. Show that $\mathfrak{A}$ is embeddable in each model of $T^0$.

7.13 Show that if a $\exists_1$-atomic theory $T$ has a model companion $T^*$, then $T^u = T^*$. 
Part II
Solutions
A
Syntax and semantics

A.1 Signature and language

In these solutions you will see various references to the compactness theorem. This is discussed in Section 1.5. You should come back to these solutions once you have read that section.

1.1 The shape of the formulas can be seen by subscripting the different uses of the variables in $\phi_3$.

$$\phi_3 := (\exists u_2)[(u_2 < v_3) \land (\exists v_2)((u_2 \equiv v_2) \land \phi_2)]$$

where

$$\phi_2 := (\exists u_1)[(u_1 < v_2) \land (\exists v_1)((u_1 \equiv v_1) \land \phi_1)]$$

where

$$\phi_1 := (\exists u_0)[(u_0 < v_1) \land (\exists v_0)((u_0 \equiv v_0) \land \phi_0)]$$

where

$$\phi_0 := (v_0 \equiv v_0)$$

Continuing this sequence in the obvious way, a simple induction gives $\partial \phi_r = \{v_r\}$ and $\partial \theta_r = \{u_r\}$ for each $r < \omega$. ■

1.2 Let $R$ be the relation symbol. It is more convenient to write this as an infix. Thus we write $uRv$ in place of $(Ruv)$. It is also convenient to sometime use $x, y, z, \ldots$ as variables.

(a) We have

(Ref) $(\forall v)[vRv]$  
(Irr) $(\forall v)[\neg(vRv)]$  
(Sym) $(\forall u, v)[uRv \rightarrow vRu]$  
(Asm) $(\forall u, v)[uRv \land vRu \rightarrow u \equiv v]$  
(Trn) $(\forall u, v, w)[uRv \land vRw \rightarrow uRw]$  
(Eqv) Ref $\land$ Sym $\land$ Trn  
(Pre) Ref $\land$ Trn  
(Pos) Pre $\land$ Asm

All of the are $\forall_1$-sentences.

In Sym and Asm we use dots in place of brackets to strengthen the implication symbol $\rightarrow$. This useful convention has now largely gone out of fashion.

(b) We work relative to the class of equivalence relations.

For each non-zero $n \in \mathbb{N}$ consider the formula

$$\land\{x_i \neq x_j \mid 1 \leq i < j \leq n\}$$

$$\land\{x_iRx_j \mid 1 \leq i < j \leq n\}$$

$$\land\{(\forall y)[x_1Ry \rightarrow \lor\{x_i \equiv y \mid 1 \leq i \leq n\}\}$$

$$\land\{(\forall y)[x_1Ry \rightarrow \lor\{x_i \equiv y \mid 1 \leq i \leq n\}\}$$

147
with \( n \) free variables \( x_1, \ldots, x_n \). Here ‘::=’ means the left hand side is defined as, and we are using a vertical display to make the formula more readable. Notice also how certain conjunctions and disjunctions are abbreviated.

A collection \( x_1, \ldots, x_n \) of the carrier of an equivalence relation satisfies
\[
\text{Block}_n(x_1, \ldots, x_n)
\]
precisely when they form a block (equivalence class) of exact size \( n \). (Notice how we are conveniently confusing elements of the structure with variables in the language.) In particular,
\[
(\exists x_1, \ldots, x_n) \text{Block}_n(x_1, \ldots, x_n)
\]
ensures there is at least one block of size \( n \).

With a slightly bit more thought we see that
\[
(\forall x_1, \ldots, x_n, y_1, \ldots, y_n) [\text{Block}_n(x_1, \ldots, x_n) \land \text{Block}_n(y_1, \ldots, y_n) \rightarrow x_1 R y_1]
\]
ensures there is no more than one block of size \( n \).

Combining these two sentences produces a
\[
\exists \land \forall
\]
sentence which ensures there is exactly one block of size \( n \).

The whole family of these sentences axiomatizes the class of equivalence relations in question. A simple application of the compactness theorem shows that we cannot replace this infinite collection by a finite collection.

An application of the compactness theorem shows that if a set of sentences can be modelled by equivalence relations with blocks of arbitrarily large finite size then the set it must also have a model with an infinite block. ■

1.3 In the first part of this solution we rather pedantically consider structures
\[
(A, *, e)
\]
where \( * \) is a binary operation written as an infix and \( e \) is a distinguished element.

(Mon) \( (\forall x, y, z)[x * (y * z) \equiv (x * y) * z] \land (\forall x)[x * e \equiv x \equiv e * x] \)

(Cmn) \( \text{Mon} \land (\forall x, y)[x * y \equiv y * x] \)

(Grp) \( \text{Mon} \land (\forall x)(\exists y)[x * y \equiv e \equiv y * x] \)

(Agp) \( \text{Grp} \land (\forall x, y)[x * y \equiv y * x] \)

We have slipped in a few abbreviations here.

To handle the class of groups we could have used an extra 1-placed operation symbol (the inverse). Technically that gives us a different notion of ‘group’ but rarely do we need to distinguish between the two. However, notice that the above is an \( \forall_2 \)-axiomatization of the notion of a group. If we use a distinguished inverse operation then we obtain an \( \forall_1 \)-axiomatization. Notice also that for the two different signatures the notion of a substructure is different. For the signature used above a substructure of a group need not be a group.
Divisible abelian groups are usually written additively (for otherwise ‘divisible’ doesn’t make sense). Let $$(A,+,0)$$ be such a group. For each element $x$ and natural number $n > 0$ we let $nx$ abbreviate $x + \cdots + x$ where the right hand side is the sum of $n$ number of $x$’s. For each such $n$ let

$$\delta_n := (\forall y)(\exists x)[nx \equiv y]$$

and then

$$\delta_1, \delta_2, \delta_3, \ldots, \delta_n, \ldots$$

are the extra axioms needed to characterize divisible abelian groups. A simple use of the compactness theorem shows that this infinite collection cannot be replaced by a finite collection.

For the next collection we think of groups written multiplicatively $$(A,\cdot,1)$$ and we hide the operation symbol.

For each $n \in \mathbb{N}$ the sentence

$$(\forall x)[x^n = 1 \rightarrow x = 1]$$

asserts that each non-neutral element (of the group) has order at least $n$. The whole collection of these sentences ensures that each non-neutral element has infinite order.

It isn’t possible to say in a first order fashion that each element of a group has finite order without putting a strict upper bound on this order. Again this follows by a simple application of the compactness theorem.

For each pair of natural numbers $m,n > 1$ consider the sentence

$$(\forall x)[x^{mn} = 1 \cdot \rightarrow x^m = 1 \lor x^n = 1]$$

and consider the whole family of these sentences. Remembering how natural numbers factorize we see that this family ensures that each element of a group has prime order. \(\blacksquare\)

1.4 We think of these structures in the form $$(A,+,0,\times,1)$$ in particular we assume that each ring is unital.

To characterize the class of rings we say that

$$(A,+,0)$$ is an abelian group \hspace{1cm} (A,\times,1)$$ is a monoid

and the two operations interact via the distributive laws. All of these are $$\forall_1$$-sentences except the one dealing with the existence of additive inverses. Furthermore, the body of each sentence is equational. (Some people might also put in

$$1 \neq 0$$
A. Syntax and semantics

but that is not wise.)

For the class of integral domains we say that the multiplication is commutative and add

\[(\forall x, y)[xy \equiv 0 \cdot \rightarrow \cdot x \equiv 0 \lor y \equiv 0]\]

which says there are no zero divisors. Observe the shape of this body. It is not just

Equations imply Equation

for there is a disjunction of equations on the right hand side. In a wider context this kind of shape can have some significance, but not here.

For the class of fields we add to the axioms for commutative rings that each non-zero element has a multiplicative inverse. This automatically prevents zero-divisors. (The status of ‘1 \neq 0’ is not clear, but almost certainly you will want to say this.)

In a field, for each \(n \in \mathbb{N}\) we may also let ‘\(n\), abbreviate

\[
1 + \cdots + 1
\]

the sum of \(n\) copies of 1. Then for each prime \(p\) the very simple sentence

\[
p \equiv 0
\]

ensures the characteristic is \(p\). The set of sentences

\[
2 \neq 0, 3 \neq 0, 5 \neq 0, \ldots, p \neq 0, \ldots
\]

ensures the characteristic is zero. By the compactness theorem this can not be replaced by a finite set of sentences.

There is no first-order way to characterize the fields with finite, non-zero, characteristic. Again this follows by the compactness theorem.

A field is algebraically closed if each non-constant polynomial over it has a root in it. For each natural number \(n > 0\) consider the sentence

\[
(\forall a_n, \ldots, a_0)[a_n \neq 0 \cdot \rightarrow \cdot (\exists x)[a_n x^n + \cdots + a_1 x + a_0 \equiv 0]]
\]

where we have used a sensible notation to make this more readable. The whole collection of these sentences ensures the field is algebraically closed. Again by the compactness theorem, this infinite set can not be replaced by a finite set of sentences.

1.5 This question has had a significant impact on the development of Mathematical Logic in general, not just Model Theory.

The original Dedekind-Peano investigation was a characterization of the structure

\[(\mathbb{N}, S, 0)\]

where \(S\) is the successor function on the natural numbers. There are a couple of simple axioms needed (saying that \(S\) is an injective function and never takes 0 as a value). The important axiom is the principle of induction.

If a subset \(X\) of \(\mathbb{N}\) contains 0 and is closed under \(S\) then it must be the whole of \(\mathbb{N}\).
This property uniquely characterizes the structure up to isomorphism. (This characterization can also be expressed in terms of category theory. The particular structure is the initial object of the category of structures of that signature. Some category theorists think this is a new observation.)

The problem with the characterization is that the induction principle as stated is not first order. So what should we do?

Peano emphasized that we should take care to note the language in which we are working. At the time set theory was being developed and there was some confusion as to what it was doing, and in some quarters complete misunderstandings. (At times Peano had to explain the difference between ‘$\in$’ and ‘$\subseteq$’, and between an object $a$ and its packaged singleton $\{a\}$. The world doesn’t change much does it!)

In the principle of induction, where should we get the set $X$ from and how should it be described? Since we are doing model theory here there is an obvious answer we should investigate. We should work in the first order language of the indicated signature.

Consider any formula

$$\phi(v, u)$$

in that language. Here $u$ is a single variable but $v$ is a batch of variables. If we fix the interpretation of the batch $v$ then we may think of the formula as naming a set. With this idea we see that

$$I(\phi) := (\forall v) [\phi(v, 0) \land (\forall u)[\phi(v, u) \rightarrow \phi(v, Su)] \cdot \rightarrow (\forall u)\phi(v, u)]$$

captures the principle of induction for a certain family of sets indexed by $v$.

We can now write down a set of axioms. A couple of trivial axioms (as mentioned above) together with all sentences $I(\phi)$ for all formulas $\phi$. (Such a collection is often called an axiom schema.) This gives us a theory formalized within the language of that signature.

Now comes the surprise.

That theory completely captures those properties of the structure that can be expressed in the language. Furthermore, the infinitely many axioms can be replaced by a small number of rather simple axioms with the same power. And, there is an effective algorithm which converts each formula of the language to an equivalent quantifier free formula relative to the axioms. Thus we can mechanically decide whether or not a sentence of the language is true or not.

Precisely how this is done is described in Section 2.3.

That result is a bit odd, isn’t it? It seems to say that number theory, or at least first order number theory, is essentially trivial.

The problem, of course, is that the first order language of that signature is so feeble that almost nothing can be expressed in it. We can’t even get at addition and multiplication.

To rectify this we try to analyse the structure

$$(\mathbb{N}, \leq, S, 0, +, \times)$$

which does look more like the basic structure for arithmetic. We now work in the first order language of that signature. We set down a few straightforward axioms, and then add all sentences $I(\phi)$ for all formulas $\phi$ of the extended language. This is what most people regard as first order arithmetic.
Now comes the crunch. This theory is incomplete. It can not even capture all the true ∀₁-sentences of the language. This is Gödel's first incompleteness theorem. Moreover, this incompleteness is not just because of an oversight; it is not because we have forgotten to include some axiom or other. No matter how we try to strengthen the theory by adding further axioms, the resulting theory will still be incomplete. That is Gödel's second incompleteness theorem.

Of course, these remarks and their ramifications ought to be explained more fully. Doing that is an interesting story involving parts of Proof Theory and Recursion Theory. Unfortunately that would take us to far from the central topic of this book. However, there are one or two more snippets of information in Section 2.3.

1.6 To be done

1.7 A metric space is a set furnished with a certain 2-placed operation to the set of real numbers. There is no way that this can be construed as a first order structure.

A topological space is a set furnished with a certain family of subsets (its topology). In the dark this does look as though it might be first order, but there are two problems. The obvious one is that the furnishings don’t belong to the standard collection of first order gadgets; it is a family of subsets. The second is that one of the conditions imposed on this family is infinitary; it must be closed under arbitrary unions.

1.8 Originally model theory had a closer relationship with the proof theory of predicate calculus. The compactness theorem was obtained as a consequence of the completeness theorem.

Now one of the family of sentences that can be proved in the predicate calculus is

\[(\forall x)\phi(x) \rightarrow (\exists x)\phi(x)\]

where \(\phi(x)\) is an arbitrary formula. This is a simple consequence of the witnessing conditions usually put into the calculus. Now consider a simple example

\[(\forall x)[x \equiv x] \rightarrow (\exists x)[x \equiv x]\]

of this kind of sentence. Here the hypothesis

\[(\forall x)[x \equiv x]\]

is trivially valid and hence the consequence

\[(\exists x)[x \equiv x]\]

is also valid. This asserts that any interpreting structure must be non-empty.
A.2 Basic notions

1.9 It is convenient to draw the posets with the comparison progressing upwards.

\[ \begin{array}{c}
\mathbb{A} \\
\mathbb{B} \\
\mathbb{C}
\end{array} \]

\[ \begin{array}{ccc}
c & c \\
b & a & b \\
a & a & d
\end{array} \]

If the letters \( a, b, c \) mean what they say then \( \mathbb{A} \) is a substructure of \( \mathbb{B} \) but not of \( \mathbb{C} \), and \( \mathbb{B} \) is not a substructure of \( \mathbb{C} \).

However, if we replace the letters by \( \bullet \) the we see that \( \mathbb{A} \) is a substructure of \( \mathbb{B} \) in two different ways, and is substructure of \( \mathbb{C} \) in four different ways. Also \( \mathbb{B} \) is a substructure of \( \mathbb{C} \) in two different ways.

1.10 Consider two relational structures

\[(A, R) \quad (B, S)\]

where each relation is 2-placed and written as an infix. The temptation is to say that an isomorphism

\[(A, R) \xrightarrow{f} (B, S)\]

is a bijection

\[f : A \rightarrow B\]

which is also a morphism, that is

\[xRy \Rightarrow f(x)Sf(y)\]

for all \( x, y \in A \). This is incorrect. This last implication must be an equivalence, that is

\[xRy \iff f(x)Sf(y)\]

for all \( x, y \in A \).

As an example consider the obvious bijective morphism between two posets.

\[ \begin{array}{c}
c \\
b \\
a
\end{array} \]

\[ \begin{array}{c}
\quad c \\
\quad b \\
\quad a
\end{array} \]

This is certainly not an isomorphism.

1.11 Most of the axioms \( \forall_2 \) and several are \( \forall_1 \)-sentences. Furthermore the body of most of these sentences are fairly simple.
In Solution 1.2 a sentence of the form
\[ \exists_2 \land \forall_2 \]
occurred. This shape usually arises when a combination of existence and uniqueness is discussed.

The most complicated sentences occurred in Solution 1.5 which discussed first order peano arithmetic. The first order induction axioms require sentences of arbitrarily high complexity. This suggests a question. What happens if we restrict the complexity of the induction axioms? The answer is known and leads to an interesting connection between a part of Proof Theory and the analysis of subrecursive hierarchies. Neither of these are central to Model Theory.

1.12 Do you really want to see these algorithms? You can not be serious!

The normal forms have some uses in theoretical analysis, but as a practical tool they are a waste of time. Even so funding bodies have wasted huge amounts of money on projects involving automated uses of these normal form. Moral: where there’s money, there is bovine fertilizer, and it often works.

A.3 Satisfaction

1.13 This kind of proof is easy but tedious. Since its the first time we have seen such a proof let’s look at the details quite closely.

We are given a structure \( \mathfrak{A} \) and a pair \( x, y \) of \( \mathfrak{A} \)-assignments. These can be fixed throughout.

To prove Lemma 1.14 consider the following property of an arbitrary term \( t \).

\[ \langle t \rangle \quad (\forall u \in \partial t)[ux = uy] \Rightarrow \mathfrak{A}[t]x = \mathfrak{A}[t]y \]

Observe the slight change of notation (\( u \) in place of \( v \)). This is done to avoid a bit of confusion later. We must show that \( \langle t \rangle \) holds for each term \( t \). We prove this by induction on the construction of \( t \). There are two base cases, a variable and a constant, and one induction step the passage across a use of an operation symbol.

For a variable \( v \) we have \( \partial v = \{ v \} \), and hence for this term \( t \) the hypothesis

\[ (\forall u \in \partial t)[ux = uy] \]

of \( \langle t \rangle \) says

\[ vx = vy \]

so that

\[ \mathfrak{A}[t]x = vx = vy = \mathfrak{A}[t]y \]

as required.

For a constant \( K \) we have \( \partial v = \emptyset \), and hence for this term \( t \) the hypothesis of \( \langle t \rangle \) is vacuously true. This is irrelevant since for this case we have

\[ \mathfrak{A}[t]x = \mathfrak{A}[K] = \mathfrak{A}[t]y \]

as required.
This deals with the two base cases.

The induction step deals with the passage across

$$t = (Ot_1 \ldots t_n)$$

where $O$ is an $n$-placed operation symbol. We have

$$\partial t = \partial t_1 \cup \cdots \partial t_n$$

so that the hypothesis of $\langle t \rangle$ gives

$$(\forall u \in \partial t_i)[ux = uy]$$

for each $1 \leq i \leq n$. Using each of the induction hypotheses

$$\langle t_1 \rangle, \ldots, \langle t_n \rangle$$

we may set

$$a_i = \mathfrak{A}[t_i]x = \mathfrak{A}[t_i]y$$

for each $1 \leq i \leq n$. This gives

$$\mathfrak{A}[t]x = \mathfrak{A}[O]a_1 \cdots a_n = \mathfrak{A}[t]y$$

as required.

This deals with the proof of Lemma 1.14. The proof of Lemma 1.16 proceeds in a similar fashion. We fix $\mathfrak{A}$ and prove

$$\langle \phi \rangle \quad (\text{for all } x, y) \left( (\forall u \in \partial \phi)[ux = uy] \implies [\mathfrak{A} \models \phi x \iff \mathfrak{A} \models \phi y] \right)$$

for an arbitrary formula $\phi$. Notice that we have now ‘un-fixed’ $x$ and $y$ and quantified over these assignments. You will see why shortly.

We proceed by induction over the construction of $\phi$ with allowable variation of $x$ and $y$.

There are four base cases two of which build on the previous result for terms. There are six induction steps. Four of these pass across a propositional connective, and are very similar to the induction step for the previous result for terms. For these cases we can pretend that $x$ and $y$ are fixed.

The two remaining induction steps pass across a use of a quantifier. Let’s look at the step across a universal quantifier.

Consider any formula $\phi$ where

$$p = (\forall v)\psi$$

for some formula $\psi$. Observe that we have

$$\partial \psi \subseteq \partial \phi \cup \{v\}$$

by the construction of $\partial \phi$. The induction hypothesis is $\langle \psi \rangle$ and we use this to verify $\langle \phi \rangle$.

Consider any pair $x, y$ of $\mathfrak{A}$-assignments, and suppose these agree on $\partial \phi$. Our problem is to verify that

$$\mathfrak{A} \models \phi x \iff \mathfrak{A} \models \phi y$$
A. Syntax and semantics

holds.

Consider any element \( a \) of \( \mathfrak{A} \). Let \( x', y' \) be the modifications of \( x, y \) obtained by replacing the \( v \)-component of each by \( a \). We now have

\[
(\forall u \in \partial \psi)[ux' = uy']
\]

and hence

\[
\mathfrak{A} \models \psi x' \iff \mathfrak{A} \models \psi y'
\]

follows by a use of the induction hypothesis.

Since the element \( a \) is arbitrary, the semantics of the universal quantifier gives (\( ? \)), as required.

1.14 (a) A simple induction shows that

\[
\mathfrak{N} \models \phi_r(a) \iff r \leq a
\]

holds for each \( a \in \mathfrak{N} \) and \( r < \omega \).

(b) This is an immediate consequence of (a).

(c) Consider any such formula \( \psi(v) \). If there is any \( a \in \mathfrak{N} \) with \( \mathfrak{N} \models \psi(a) \) then, by (a), we have \( r \leq a \) for each \( r < \omega \). This is impossible, and hence there is no such \( a \). Thus \( \mathfrak{N} \models \neg(\exists v)\psi(v) \) must hold.

1.15 (a) For an arbitrary \( n \in \mathbb{N} \) with \( n \geq 2 \) let \( \sigma_n \) be

\[
(\exists x_1, \ldots, x_n)[\bigwedge \{ x_i \neq x_j \mid 1 \leq i \leq j \leq n\}]
\]

to obtain an \( \exists_1 \)-sentence that does the required job. (With a bit of give and take we see that this also works for \( n = 0, 1 \).)

For each \( n \) the \( \exists_1 \land \forall_1 \)-sentence

\[
\tau_n \::= \sigma_n \land \neg\sigma_{n+1}
\]

will do.

(b) The set

\[
Inf = \{ \sigma_n \mid n \in \mathbb{N} \}
\]

will do. A use of the compactness theorem shows that no finite set of sentences will do this job.

(c) A similar use of the compactness theorem shows that there is no such set \( Fin \). ■

1.16 Let \( f \) be the 1-placed operation symbol. We say that \( f \) is injective but not surjective (from the carrier to the carrier). If the carrier is finite then any injective endo-function is also surjective. Thus for \( \sigma \) we take the conjunction of two sentences.

\[
(\forall x, y)[fx \equiv fy \rightarrow x \equiv y] \quad (\exists y)(\forall x)[fx \neq y]
\]

This gives an \( \exists_2 \)-sentence.

For each \( n > 0 \) let \( \sigma_n \) be any sentence such that

\[
\mathfrak{A} \models \sigma_n \iff \mathfrak{A} \text{ has size at least } n
\]
for each structure $\mathfrak{A}$. Such sentences can be produced using only the equality symbol. Let

$$\Sigma = \{ \sigma_n \mid 0 < n \in \mathbb{N} \}$$

so that

$$\mathfrak{A} \models \Sigma \iff \mathfrak{A} \text{ is infinite}$$

for each structure $\mathfrak{A}$. We show that this set $\Sigma$ can not be replaced by a single sentence.

By way of contradiction suppose there is a sentences $\tau$ such that

$$\mathfrak{A} \models \tau \iff \mathfrak{A} \text{ is infinite}$$

for each structure $\mathfrak{A}$. Then

$$\mathfrak{A} \models \neg \tau \iff \mathfrak{A} \text{ is finite}$$

for each structure $\mathfrak{A}$. In particular

$$\Sigma \cup \{ \neg \tau \}$$

is finitely satifiable, and hence satisfiable by the compactness theorem. Any model of this set must be both infinite and finite, which is impossible.

1.17 The first part is another tedious proof by induction over the construction of the quantifier-free formula $\delta$. There are four base cases and four induction steps. Most of the arguments are similar to those used in Solution 1.13.

To deal with the two ‘non-trivial’ base cases we must first check that

$$\mathfrak{A}[t]x = \mathfrak{B}[t]x$$

for each term $t$ and $\mathfrak{A}$-assignment $x$. The proof of this proceeds by induction over the construction of $t$. The base cases are essentially the definition of the substructure relation.

The second part is slightly more interesting. Suppose

$$\mathfrak{A} \models \theta(a)$$

for some $\exists_1$-formula $\theta(v)$ and point $a$ which matches the batch $v$ of variables. We know that $\theta$ is

$$(\exists u)\delta(v, u)$$

for some quantifier-free formula $\delta$ and appropriate batch $u$ of variables. (Notice the useful abbreviation $(\exists u)$ here.) from the above supposition we have

$$\mathfrak{A} \models \delta(a, a')$$

for some point $a'$ which matches the batch $u$. Since $\delta$ is quantifier-free he first part gives

$$\mathfrak{B} \models \delta(a, a')$$

and hence

$$\mathfrak{B} \models \theta(a)$$

as required.

This implication can not be strengthened to an equivalence. To see this suppose $\mathfrak{A} \subseteq \mathfrak{B}$ where both are finite with $\mathfrak{A}$ small but $\mathfrak{B}$ large. Using Solution 1.15 we have

$$\mathfrak{A} \models \neg \sigma_n \quad \mathfrak{B} \models \sigma_n$$

for some $n$. 

$\blacksquare$
A. Syntax and semantics

A.4 Consequence

1.18 Much of this is done in Solution 1.2, so we need not repeat those details.

(b) For $m \in \mathbb{N}$ we easily write down

$$\text{For each } x_1, \ldots, x_{m+1} \text{ two of them are equivalent}$$

as an $\forall_1$-sentence. This says that the structure has no more than $m$ blocks.

For $n \in \mathbb{N}$ we can write down

$$\text{For each } x_1, \ldots, x_{n+1} \text{ if these are pairwise equivalent then two of them are equal}$$

as a slightly more complicated $\forall_1$-sentence. This says that each block has no more then $n$ members.

(c) [This can’t be done -- need to sort out]

(d) We may use the $\forall_1$-formulas

$$\text{Block}_n(x_1, \ldots, x_n)$$

of Solution 1.2. With these the whole family of $(\forall_2 \lor \exists_2)$-sentences

$$(\exists x_1, \ldots, x_n)\text{Block}_n(x_1, \ldots, x_n) \cdot \to \cdot (\exists x_1, \ldots, x_n, x_{n+1})\text{Block}_n(x_1, \ldots, x_n, x_{n+1})$$

will do. ■

1.19 (a) A linear ordering is dense if for each pair $x < z$ of elements there is something strictly in between $x < y < z$.

A linear ordering is discrete if for each pair $x < z$ of elements there is a pair $x \leq u < w \leq z$ with nothing in between, no $u < v < w$.

You have to be a bit careful with both of these axiomatization because the linear order may have an end point at one extreme or the other.

(b) The formulas $\theta$ and $\psi$ shouldn’t cause any trouble.

For the formula $\phi$ we just have to be patient and a bit careful. ■

1.20 Apart from torsion-free abelian groups, these classes are axiomatized in Solution 1.3. An abelian group is torsion-free if, in additive notation, we have

$$na = 0 \implies a = 0$$

for each element $a$ and $n \in \mathbb{N}$.

The main problem with the formulas $\theta, \psi, \phi$ is remembering what the group theoretic notions are. We use the multiplicative notation.

An element $a$ is a commutator if

$$a = xyx^{-1}y^{-1}$$

for some elements $x, y$. This description makes use of the inverse operation, but that can be eliminated by a couple of existential quantifiers.

An element $a$ is in the commutator of the element $b$ precisely when $ab = ba$. 
An inner automorphism has the form
\[ a \mapsto xax^{-1} \]
for some element \( x \). So \( a \) is taken to \( b \) by some inner automorphism if
\[ xa = bx \]
for some element \( x \). ■

1.21 *Dealt with earlier -- sort out*

1.22 The axioms for a field and the axioms for a linearly ordered set are easy, and give \( \forall_2 \)-sentences.

There are various axioms concerning the interaction of the field operations and the linear order. For instance, for each element \( x \) precisely one of
\[ x < 0 \quad x = 0 \quad 0 < x \]
holds. The ordering is given by the positive cone, that is
\[ x \leq y \iff (y - x) \geq 0 \]
for each \( x, y \). All of these axioms are \( \forall_1 \)-sentences (and it is easy to forget some of them).

Next we have to say that the field is ‘algebraically closed as far as it can be’. Firstly we need
\[ (\forall x)(\exists y)[x = y^2 \lor x = -y^2] \]
to give all the square roots the field should have. Also, each polynomial of odd degree must have a root. Thus we require
\[ (\forall a_0, \ldots, a_{2n})(\exists x)[a_0 + a_1 x + \cdots + a_{2n} x^{2n} + x^{2n+1} = 0] \]
for each \( n \in \mathbb{N} \).

This infinite set of \( \forall_2 \)-sentences form the axioms for a real closed field. Of course, these don’t characterize \( \mathbb{R} \).

The crucial missing property is the completeness of \( \mathbb{R} \). This can be captured in one of two ways.

We may say that the field is Dedekind complete; that each bounded subset has a least upper bound.

Or we may say that the field is Cauchy complete; that each cauchy convergent sequence has a limit.

Neither of these are first order properties (and in some non-standard universes are not equivalent).

There has been a lot of model theoretic analysis of real closed fields and no-standard extension of \( \mathbb{R} \). However, this is beyond the scope of this book. ■

1.23 We have to be careful here for we are not allowed to quantifier over the ring \( R \).

There are several axioms but the two crucial ones are
\[ a(r + s) = ar + as \quad (ar)s = a(rs) \]
for each element \( a \) of the module and elements \( r, s \) of the ring. There are two additions here, one on the module and one on the ring. Let us write \(+\) for the addition on the module. With this notation the above two axioms become

\[
(\forall v)[f_{r+s}v \simeq (f_r v + f_s v)] \quad (\forall v)[f_s(f_r v) \simeq f_{rs}v]
\]
in more or less the official language. Remember that here the quantified variable \( v \) ranges over the module.

For a left module the two crucial axioms are

\[
(s + r)a = sa + ra \quad \text{and} \quad s(ra) = (sr)a
\]
respectively. These become

\[
(\forall v)[f_{s+r}v \simeq (f_s v + f_r v)] \quad (\forall v)[f_s(f_r v) \simeq f_{rs}v]
\]
respectively.

This is one subject where the official format of a first order language doesn’t help. ■

1.24 (i) \( \Rightarrow \) (ii). Suppose \( T \cup \{\sigma\} \) is consistent and consider any \( \mathfrak{A} \models T \cup \{\sigma\} \). Then, using (i), we have

\[
\mathfrak{B} \models T \implies \mathfrak{B} \equiv \mathfrak{A} \implies \mathfrak{B} \models \sigma
\]
so that \( \sigma \in T \).

(ii) \( \Rightarrow \) (iii). Suppose \( T \subseteq T' \) where \( T' \) is consistent. Then (ii) gives

\[
\sigma \in T' \implies T \cup \{\sigma\} \text{ consistent} \implies \sigma \in T
\]
and hence \( T = T' \).

(iii) \( \Rightarrow \) (iv). We verify the contrapositive of (iv). Consider a pair of sentences \( \sigma, \tau \) neither of which is in \( T \). Then both

\[
T \cup \{\neg \sigma\} \quad T \cup \{\neg \tau\}
\]
are consistent, and hence (iii) (or even (ii)) gives

\[
T \vdash \neg \sigma \quad T \vdash \neg \tau
\]
so that \( \neg \sigma \land \neg \tau \in T \), which leads to \( \sigma \lor \tau \notin T \).

(iv) \( \Rightarrow \) (v). For each sentence \( \sigma \), we have \( \sigma \lor \neg \sigma \in T \).

(v) \( \Rightarrow \) (vi). Consider any model \( \mathfrak{A} \) of \( T \). We have \( T \subseteq Th(\mathfrak{A}) \). If these two are not equal then there is some sentence \( \sigma \notin T \) with \( \mathfrak{A} \models \sigma \). But then (v) gives \( \neg \sigma \in T \), and hence \( \mathfrak{A} \models \neg \sigma \), which is a contradiction.

(vi) \( \Rightarrow \) (i). We have \( T = Th(\mathfrak{A}) \) for some model \( \mathfrak{A} \). Consider any other model \( \mathfrak{B} \). For each sentence \( \sigma \) we have

\[
\mathfrak{A} \models \sigma \implies \sigma \in T \implies \mathfrak{B} \models \sigma
\]
to give \( \mathfrak{A} \equiv \mathfrak{B} \). For each pair of models \( \mathfrak{B}, \mathfrak{C} \) we have

\[
\mathfrak{B} \equiv \mathfrak{A} \equiv \mathfrak{C}
\]
to give $\mathcal{B} \equiv \mathcal{C}$, and hence $T$ is complete. 

1.25 Let $\Psi$ be the full binary splitting tree (sometimes called the Cantor tree). The nodes $\mu, \nu, \ldots$ of $\Psi$ are the finite lists 

$$i(0)i(1) \cdots i(n-1)$$

taken from $\{0, 1\}$. This particular list has length $n$, and the empty list

$$\bot$$

of length 0 is allowed. These nodes are partially ordered by extension.

$$
\begin{array}{cccc}
\vdots & \vdots & \vdots & \vdots \\
000 & 001 & 010 & 011 \\
00 & 01 & 10 & 11 \\
0 & 1 & 0 & 1 \\
\bot & \bot & \bot & \bot \\
\end{array}
$$

In particular

$$\nu \leq \mu, \mu \mid \nu$$

is a typical node $\nu$ and its two successors $\nu_0, \nu_1$.

We write

$$\nu \leq \mu \quad \mu \mid \nu$$

to indicate, respectively, that $\mu$ extends $\nu$, and that $\mu$ and $\nu$ are incomparable.

A branch of this tree is just a function

$$p : \mathbb{N} \longrightarrow \{0, 1\}$$

and there are $2^{\aleph_0}$ such branches. This branch lies above the nodes

$$\bot, p(0), p(0)p(1), p(0)p(1)p(2), \ldots$$

the finite initial sections of $p$.

Assuming that no finite extension of the theory $T$ is complete, we produce a family

$$\Sigma = \{\sigma_{\nu} \mid \nu \in \Psi \bot\}$$

of sentences, each consistent with $T$, with

$$\sigma_{\bot} = \text{true}$$

and such that

$$(i) \quad \nu \leq \mu \implies T \vdash \sigma_{\mu} \rightarrow \sigma_{\nu} \quad (ii) \quad \mu \mid \nu \implies T \vdash \neg \sigma_{\mu} \lor \neg \sigma_{\nu}$$

hold for all nodes $\mu, \nu$ of $\Psi$. 
We generate this family by recursion up the tree $\Psi$. Suppose we have generated $\sigma_\nu$ for some $\nu \in \Psi$. Thus $T \cup \{\sigma_\nu\}$ is consistent, and does not axiomatize a complete extension of $T$. Thus there is a sentences $\tau$ such that both

$$T \cup \{\sigma_\nu, \tau\} \quad T \cup \{\sigma_\nu, \neg \tau\}$$

are consistent. We set

$$\sigma_{\nu 0} = \sigma_\nu \land \tau \quad \sigma_{\nu 1} = \sigma_\nu \land \neg \tau$$

to produce the two immediate successors of $\sigma_\nu$. Note that

$$T \vdash \sigma_{\nu i} \rightarrow \sigma_\nu \quad T \vdash \neg \sigma_{\nu 0} \lor \neg \sigma_{\nu 1}$$

hold.

Once we have the family $\Sigma$ we set

$$\Sigma_p = \{\sigma_\nu \mid \nu < p\}$$

for each branch $p$ of $\Psi$. By (i) we see that each such theory $T \cup \Sigma_p$ is consistent. Let

$$\mathfrak{A}_p \models T \cup \Sigma_p$$

to produce a complete extension $T(p) = Th(\mathfrak{A}_p)$ of $T$. We show that these are pairwise distinct.

Consider distinct branches $p, q$ of $\Psi$. There are nodes $\mu < p, \nu < q$ with $\mu | \nu$. Then (ii) gives

$$\mathfrak{A}_p \models \neg \sigma_\nu \quad \mathfrak{A}_q \models \neg \sigma_\mu$$

so that $\mathfrak{A}_\mu \not\equiv \mathfrak{A}_q$, as required.

1.26 (a) For each $a \in A$ we have

$$s = \iota(a) \leq \iota(a)$$

and hence the equivalence gives

$$a \leq \iota(s) = (\iota \circ \iota)(a)$$

to show that $\iota \circ \iota$ is inflationary. A similar argument shows that $\iota \circ \iota \circ \iota$ is inflationary.

(b) For each $a \in A$ we have

$$\iota(a) \leq (\iota \circ \iota \circ \iota)(a)$$

since $\iota \circ \iota$ is inflationary. Using the inflationary property twice we have

$$a \leq (\iota \circ \iota \circ \iota \circ \iota)(a)$$

and hence

$$(\iota \circ \iota \circ \iota)(a) \leq \iota(a)$$

by the given equivalence. This shows that

$$\iota \circ \iota \circ \iota = \iota$$
and the other equality follows by a similar argument.

(c) Suppose that
\[ a \leq b \]
for some \( a, b \in A \). Then
\[ a \leq b \leq (\uparrow \circ \uparrow)(b) \]
so that
\[ \uparrow(b) \leq \uparrow(a) \]
by the given equivalence. This shows that \( \uparrow \) is antitone, and a similar argument shows that \( \downarrow \) is antitone.

(d) By (a) both composites are inflationary. By two uses of (c) both composites are monotone. By two uses of (b) both composites are idempotent. \( \blacksquare \)

A.5 Compactness

1.27 (a) If \( \Sigma \vdash \tau \), then \( \Sigma \cup \{ \neg \tau \} \) does not have a model, and hence, by compactness, some finite part of this set does not have a model. In other words, there is some finite \( \Gamma \subseteq \Sigma \) such that \( \Gamma \cup \{ \neg \tau \} \) does not have a model. With this \( \Gamma \) we have \( \Gamma \vdash \tau \), to give the required result.

(b) Suppose \( \Sigma \) has the closure properties. We show that
\[ \Sigma \vdash \tau \implies \tau \in \Sigma \]
holds, and hence \( \Sigma \) is a theory. (The other required implication is trivial.)

Suppose \( \Sigma \vdash \tau \). By part (a) we have \( \Gamma \vdash \tau \) for some finite part
\[ \Gamma = \{ \sigma_m, \ldots, \sigma_1 \} \]
of \( \Sigma \). But now the sentence
\[ \sigma_m \rightarrow \cdots \sigma_1 \rightarrow \tau \]
is logically valid, and hence belongs to \( \Sigma \). With this we may strip of \( \sigma_m, \ldots, \sigma_1 \) in turn, to get \( \tau \in \Sigma \). \( \blacksquare \)

1.28 Since each \( K_r \) is strictly elementary we have \( K_r = M(\sigma_r) \) for some family
\[ \Sigma = \{ \sigma_r \mid r < \omega \} \]
of sentences. Since \( K_r \supseteq K_{r+1} \) we may assume that \( \vdash \sigma_{r+1} \rightarrow \sigma_r \) holds. (This may be achieved by taking conjunctions of the original \( \sigma_r \).)

We see that \( K = M(\Sigma) \), so that \( K \) is elementary.

By way of contradiction, suppose that \( K \) is strictly elementary. Thus \( K = M(\tau) \) for some sentence \( \tau \). In particular
\[ \Sigma \vdash \tau \]
and
\[ \vdash \tau \rightarrow \sigma_r \]
for each \( r < \omega \). These give some particular \( n < \omega \) such that \( \models \tau \leftrightarrow \sigma_n \), and hence \( K = M(\sigma_n) = K_n \), which is not so. 

1.29 Let 
\[
K = M(\sigma) \quad L = M(L) \quad R = M(R)
\]
where \( \sigma \) is a sentence and \( L, R \) are sets of sentences. Since \( L \cap R = \emptyset \), we see that \( L \cup R \) does not have a model, and hence compactness gives sentences \( \lambda, \rho \) such that
\[
\models \lambda \land \sigma \quad \models \rho \land \sigma \quad \models \neg \lambda \lor \neg \rho
\]
hold. We show that, in fact,
\[
L = M(\lambda \land \sigma) \quad R = M(\rho \land \sigma)
\]
hold.

Consider any \( \mathfrak{A} \in M(\lambda \land \sigma) \). Since \( \mathfrak{A} \models \sigma \), we have \( \mathfrak{A} \in K = L \cup R \). Since \( \mathfrak{A} \models \lambda \), we have \( \mathfrak{A} \models \neg \rho \), and hence \( \mathfrak{A} \notin R \), to give \( \mathfrak{A} \in L \), as required.

1.30 (a,b) Let \( \alpha \) be conjunction of the usual axioms for fields. This shows that \( F \) is strictly elementary, and axiomatizes \( T \).

For each prime \( p \) let
\[
(p \equiv 0) \quad \text{abbreviate} \quad 1 + \cdots + 1 \equiv 0
\]
where there are \( p \) number of 1s in this sum. Then
\[
\alpha \land (p \equiv 0)
\]
axiomatizes \( T_p \), to show that \( F_p \) is strictly elementary. Also
\[
\{\alpha\} \cup \{(p \not\equiv 0) \mid p \text{ a prime}\}
\]
axiomatizes \( T_0 \), to show that \( F_0 \) is elementary.

We have
\[
F_0 \cup F_f = F \quad F_0 \cap F_f = \emptyset
\]
so that, by Exercise 1.29, if \( F_f \) is elementary then \( F_0 \) is strictly elementary. Part (c) show that this is not the case.

(c) For each \( \tau \in T_0 \) we have
\[
T \cup \{(p \not\equiv 0) \mid p \text{ a prime}\} \models \tau
\]
and hence
\[
T \cup \{(2 \not\equiv 0), \ldots, (p \not\equiv 0)\} \models \tau
\]
for some sufficiently large prime \( p \). Any field of larger prime characteristic show that this can not be so.

In particular, \( T_0 \) is not finitely axiomatizable, and \( F_0 \) is not strictly elementary.

(d) The same method show that
\[
T_f \cup \{(p \not\equiv 0) \mid p \text{ a prime}\}
\]
is finitely satisfiable, and hence has a model.

1.31 For the first part all we need to observe is that $Th(\mathcal{R})$ can be formalised in a countable language.

The second part depends on which signature is used.

If we view $\mathcal{R}$ as a line $(\mathbb{R}, \leq)$ (in the sense of subsection 2.2) then the rational line $(\mathbb{Q}, \leq)$ is the only possibility.

If we view $\mathcal{R}$ in the most natural way as a linearly ordered real field, then the real part of the algebraic closure of $\mathbb{Q}$ will do (and there are other possibilities).

This second part requires some results not yet covered and some information about the structure of real fields.

1.32 Well, go on!
The effective elimination of quantifiers

B.1 The generalities of quantifier elimination

Since there are no operation symbols in the signature, each atomic sentence has one of the shapes

\[ \text{true} \quad \text{false} \quad (K_1 \cong K_2) \quad RK_1, \ldots, K_m \]

where \( K_1, \ldots, K_m \) are constant symbols and \( R \) is an \( m \)-placed relation symbol. Since the signature is finite, there are only finitely many such atomic sentences. Let \( \alpha_1, \ldots, \alpha_n \) be these atomic sentences.

Each quantifier-free sentence is \( T \)-equivalent to a conjunction of disjunction of literals. Each such disjunction can be put in the form

\[ \pm \alpha_1 \lor \cdots \lor \pm \alpha_n \]

where each component \( \pm \alpha \) is \( \alpha \) or \( \neg \alpha \), as appropriate. There are only \( 2^n \) possible disjunctions of this kind. Each quantifier-free sentence is \( T \)-equivalent to a conjunction of such disjunctions, and hence there are only finitely many quantifier-free sentences up to \( T \)-equivalence.

But every sentence is \( T \)-equivalent to a quantifier-free sentence, and hence the boolean algebra of sentences modulo \( T \) is finite.

Each complete extension is determined by the set of sentences that need to be added to \( T \). Each such set is finite and, in fact, bounded by the size of the boolean algebra of all sentences. Thus \( T \) has only finitely many complete extension.

B.2 Linear orders

Most of the axioms are \( \forall_1 \) but the important ones are \( \forall_2 \). For instance

\[ (\forall v)(\exists u)[u < v] \quad (\forall u, w)[(u < v) \rightarrow (\exists v)[u < v < w]] \quad (\forall v)(\exists w)[v < w] \]

say there is no first point, the line is dense, and there is no last point, respectively.

For the proof of Lemma 2.5 we work in the extended language with both \( \leq \) and \( < \) as relation symbols.

Each \( QF \)-formula is built up from atomic formulas using \( \neg, \wedge, \lor \). Each use of \( \neg \) can be driven deeper using the de Morgan laws. Any occuring double negation can be omitted. Any negated atomic formula can be replaced by a negation-free equivalent. Thus we obtain an equivalent formula built up using only \( \wedge \) and \( \lor \). If required this can be put into disjunctive or conjunctive normal form.

2.4 Suppose

\[ (\exists w)[(u_1 < w) \land \cdots \land (u_i < w) \land (w < v_1) \land \cdots \land (w < u_r)] \]
is the formula $\theta$. This is not the most general possible shape for $\theta$, but it is the one that contains all the complications that have to be handled. For this $\theta$ the conjunction

$$\bigwedge \{u_i < v_j \mid 1 \leq i \leq l, 1 \leq j \leq r\}$$

is the corresponding $QF$-formula $\gamma$.

From $\theta$ we have some $w$ with

$$u_i < w < v_j$$

for each relevant $i, j$. Thus $u_i < v_j$. This verifies

$$T \vdash \theta \rightarrow \gamma$$

and we now require the converse.

Consider any line $A$ in which $\gamma$ holds for some $u_1, \ldots, u_l, v_1, \ldots, v_r$. Let

$$u = \max \{u_i \mid 1 \leq i \leq l\} \quad v = \min \{v_j \mid 1 \leq j \leq r\}$$

so that

$$u < v$$

by $\gamma$. Since the line is dense there is some element $w$ with

$$u < w < v$$

and this provides a witness for $\theta$. ■

2.5 Let $T$ be the theory of lines and consider any sentence $\sigma$. Then

$$T \vdash \sigma \leftrightarrow \delta$$

for some quantifier-free sentence $\delta$. But, in this language, the only two quantifier-free sentences are true and false. Thus one of

$$T \vdash \sigma \quad T \vdash \neg \sigma$$

must hold. This shows that $T$ is complete. ■

2.6 In some ways this is easier. For the algorithm corresponding to that in the proof of Theorem 2.6 there are fewer cases to consider. ■

2.7 Both $\mathbb{Q}$ and $\mathbb{R}$ are models of the theory $T$ (of lines), which is complete, so that $Th(\mathbb{Q}) = T = Th(\mathbb{R})$ and hence $\mathbb{Q} \equiv \mathbb{R}$.

This shows that the (Dedekind) completeness of the line $\mathbb{R}$ can not be captured in a first order way (at least within the languages of lines). ■

2.8 Let $(A, \leq)$ be a linearly ordered set. We form a new linear ordering by replacing each element of $A$ by a copy of $\mathbb{Q}$. These blocks are ordered as the occur in $A$. This gives the ordinal product $\mathbb{Q} \cdot A$.

In more detail, we look at the set of pairs

$$(p, a)$$
where $p \in \mathbb{Q}$ and $a \in A$. We compare these by

$$(p, a) \leq (q, b) \iff (a < b) \text{ or } [(a = b) \text{ and } (p \leq q)]$$

(for $p, q \in \mathbb{Q}$ and $a, b \in A$). It is routine to check that this gives a line, and $a \mapsto (0, a)$ is an embedding.

2.9 The $\aleph_0$-categoricity is the simplest application of the back-and-forth method. This is discussed in section 5.1.

Consider any cardinal $\kappa > \aleph_0$. There are several methods of producing non-isomorphic lines of cardinality $\kappa$. Here is one of them.

Consider any limit ordinal $\lambda$ of cardinality $\kappa$. Let $\lambda^*$ be the reverse of $\lambda$, and consider the ordinal products

$$\mathbb{Q} \cdot \lambda \quad \mathbb{Q} \cdot \lambda^*$$

both of which are lines of cardinality $\kappa$. We know that, as a linear order, $\lambda$ is embedded in the first, and $\lambda^*$ is embedded in the second. We show that $\lambda^*$ cannot be embedded in the first, and so the two are not isomorphic.

By way of contradiction suppose $\lambda^*$ can be embedded in $\mathbb{Q} \cdot \lambda$. Thus we have an indexed family

$$a(i) = (p(i), \alpha(i)) \quad (i < \lambda)$$

of elements of the line such that

$$j < i < \lambda \implies a(i) < a(j)$$

holds. Remembering how these pairs are compared, we see that

$$j < i < \lambda \implies \alpha(i) \leq \alpha(j)$$

holds. In other words, we have a descending chain through $\lambda$. But $\lambda$ is an ordinal, so this chain must stabilize after a finite number of steps. In other words, there is some finite $k < \omega$ such that

$$k \leq j < i < \lambda \implies \alpha(i) = \alpha(j)$$

and hence

$$k \leq j < i < \lambda \implies p(i) < p(j)$$

holds. This gives a subset of $\mathbb{Q}$ of cardinality $|\lambda| = \kappa$, which is impossible.

B.3 The natural numbers

2.10 We are required to show

(a) $T^+ \vdash (2)$
(b) $T^+ \vdash (3)$

using the labelling of Definition 2.10.

Handling internal (formal) induction can be quite tricky, especially when it is combined with external induction.

Let $\phi(v)$ be any formula with just the free variable $v$, as indicated. Notice that

(c) $\phi(0) \land (\forall v)\phi(Sv) \rightarrow (\forall v)\phi(v)$
is a simple consequence of the induction axiom for \( \phi \). This is the parameter-free case induction.

(a) Let
\[
\phi(v) := (v \equiv 0) \lor (\exists w)[Sw \equiv v]
\]
so that
\[
T^+ \vdash \phi(0) \quad T^+ \vdash \phi(Sv)
\]
(for we can witness the appropriate disjunct of \( \phi \)). The required result
\[
T^+ \vdash (\forall v)\phi(v)
\]
is an immediate application of (case).

(b) Let
\[
\phi_k(v) := (S^{k+1}v \neq v)
\]
(for each \( k \in \mathbb{N} \)). First we check that
\[
(i) \quad T^+ \vdash \phi_0(v) \quad (ii) \quad T^+ \vdash \phi_{k+1}(0) \quad (iii) \quad T^+ \vdash \phi_k(v) \rightarrow \phi_{k+1}(Sv)
\]
hold.

We have
\[
T^+ \vdash \phi_0(0)
\]
by axiom (0), and
\[
T^+ \vdash (S(Sv) \equiv Sv) \rightarrow (Sv \equiv v)
\]
by axiom (1), so that
\[
T^+ \vdash \phi_0(v) \rightarrow \phi_0(Sv)
\]
and hence a use of internal induction gives (i).

An instance of axiom (0) fives (ii).

By axiom (1) we have
\[
T^+ \vdash (S^{k+2}v \equiv Sv) \rightarrow (S^{k+1}v \equiv v)
\]
which gives (iii).

We now show
\[
[k] \quad T^+ \vdash (\forall v)\phi_k(v)
\]
by an external induction over \( k \).

The base case [0], is just (i).

For the induction step, \([k] \Rightarrow [k + 1]\), we have
\[
T^+ \vdash (\forall v)\phi_k(v) \rightarrow (\forall v)\phi_{k+1}(Sv)
\]
by (iii), so that \([k]\), (ii), and (case) give \([k + 1]\).

\[\blacksquare\]

2.11 We prove the two required implications separately.

For each \( l \leq k \) we have a sequence of implications
\[
T \vdash (S^{k+1}w \equiv v) \land (v \equiv S^l0) \rightarrow (S^{k+1}w \equiv S^l0) \rightarrow (S^{(k-l)+1}w \equiv 0) \rightarrow \text{false}
\]
by the properties of equality, axiom (2), and axiom (0). Thus
\[ T \vdash (\exists w)[S^{k+1}w \equiv v] \rightarrow (v \not\equiv \top) \]
to give
\[ T \vdash (\exists w)[S^{k+1}w \equiv v] \rightarrow (v \not\equiv \top) \land \cdots \land (v \not\equiv \top) \]
which is the first implication.

For the converse let \( \phi_k(v) := (v \equiv \top) \lor \cdots \lor (v \equiv \top) \lor (\exists w)[S^{k+1}w \equiv v] \)
(for each \( k \in \mathbb{N} \)). We first observe that
\[ (iv) \ T \vdash \phi_{k+1}(0) \quad (v) \ T \vdash \phi_{k+1}(Sv) \leftrightarrow \phi_k(v) \]
hold. The first, \((iv)\), is trivial. For the second, \((v)\), we have
\[ T \vdash \phi_{k+1}(Sv) \leftrightarrow (Sv \equiv \top) \lor (Sv \equiv \top) \lor \cdots \lor (Sv \equiv \top) \lor (\exists w)[S^{k+2}w \equiv Sv] \]
\[ \leftrightarrow (Sv \equiv \top) \lor \cdots \lor (Sv \equiv \top) \lor (\exists w)[S^{k+2}w \equiv Sv] \]
\[ \leftrightarrow \phi_k(v) \]
using first axiom (0), and then axiom (1). With these we show
\[ [k] \ T \vdash (\forall v)\phi_k(v) \]
by an external induction over \( k \).

The base case \([0]\), is axiom (2).

For the induction step, \([k] \Rightarrow [k + 1]\), we have
\[ T \vdash (\forall v)\phi_{k+1}(Sv) \]
by \([k]\) and \((v)\), so that \((iv)\) and \((\text{case})\) give \([k + 1]\). \(\blacksquare\)

2.12 (a) Consider any element \( a \) of any model \( \mathfrak{A} \) of \( T \). We may generate a chain
\[ \{a(r) \mid r \in \mathbb{N}\} \]
of elements of \( \mathfrak{A} \) by setting \( a(r) = S^r a \). By axiom (3) (as given in Solution 2.11), all these elements are distinct. Thus we generate a copy of \( \mathfrak{N} \). In particular, when \( a \) is the distinguished element of \( \mathfrak{A} \), we identify this copy with \( \mathfrak{N} \). Recall that axiom (0) says that 0 does not have a predecessor.

Now consider any element \( a \) not in (this canonical copy of) \( \mathfrak{N} \). Using axiom (2) we may regress to form a chain
\[ \{a(-r) \mid r \in \mathbb{N}\} \]
of elements of \( \mathfrak{A} \) where \( a(0) = a \) (as before) and \( Sa(-r + 1) = a(-r) \) for each \( r \in \mathbb{N} \).
Using axiom (3) we see these two chains form a copy of \( \mathfrak{Z} = (\mathbb{Z}, S, 0) \).

We now remove \( \mathfrak{N} \) from \( \mathfrak{A} \). What we have left can be partitioned into blocks each of which is a copy of \( \mathfrak{Z} \). We count the number of such blocks, to see that
\[ \mathfrak{A} = \mathfrak{N} \cup \mathfrak{Z}^{(n)} \]
where the superscript indicates there are \( \kappa \) copies of \( Z \). In particular, this \( \kappa \) completely determines the (isomorphism) type of \( \mathfrak{A} \). Let us write \( \mathfrak{A}(\kappa) \) for this model. Thus every model is isomorphic to \( \mathfrak{A}(\kappa) \) for precisely one \( \kappa \).

Notice that the cardinality of \( \mathfrak{A}(\kappa) \) is \( \kappa \) or \( \aleph_0 \), whichever is the larger.

(b) For uncountable \( \kappa \), each model of \( T \) of cardinality \( \kappa \) is isomorphic to \( \mathfrak{A}(\kappa) \).

c) Each countable model of \( T \) is isomorphic to precisely one of

\[
\mathfrak{N} = \mathfrak{A}(0) \subseteq \mathfrak{A}(1) \subseteq \cdots \subseteq \mathfrak{A}(r) \subseteq \cdots \subseteq \mathfrak{A}(\aleph_0)
\]

which is a chain of order type \( \omega + 1 \).

This spectrum of models is typical of a large class of theories.

[What are this kind of theories called?]

2.13 (a) Since \( \mathfrak{N} \models T \), we have

\[
T \vdash \delta \iff \mathfrak{N} \models \delta
\]

(for each quantifier-free sentence), so it suffices to show that converse. There are two ways to do this.

Each atomic sentence \( \alpha \) of this language has the shape

\[
\left( ^m \equiv ^n \right)
\]

for some \( m, n \in \mathbb{N} \) (or is one of true or false). Then

\[
\mathfrak{N} \models \alpha \implies (m = n) \implies T \vdash \alpha
\]

\[
\mathfrak{N} \models \neg \alpha \implies (m \neq n) \implies T \vdash \neg \alpha
\]

where the last implication makes use of axioms (1) and (0). Using this we show

\[
\mathfrak{N} \models \delta \implies T \vdash \delta \quad \mathfrak{N} \models \neg \delta \implies T \vdash \neg \delta
\]

for each quantifier-free sentence \( \delta \). We proceed by induction on the construction of \( \delta \). Notice that we verify the two implications in tandem.

The base case, where \( \delta \) is atomic, is just the observation above. For the induction step we survey the possible shapes of \( \delta \). There are four cases

\[
\begin{align*}
\delta &= (\alpha \land \beta) & T \vdash \neg \delta & \iff (\neg \alpha \lor \neg \beta) \\
\delta &= (\alpha \lor \beta) & T \vdash \neg \delta & \iff (\neg \alpha \land \neg \beta) \\
\delta &= (\alpha \to \beta) & T \vdash \neg \delta & \iff (\alpha \land \neg \beta) \\
\delta &= \neg \gamma & T \vdash \neg \delta & \iff \gamma
\end{align*}
\]

where \( \alpha, \beta, \gamma \) are simpler sentences.

For the first case we have

\[
\begin{align*}
\mathfrak{N} \models \delta & \implies \mathfrak{N} \models \alpha \text{ and } \mathfrak{N} \models \beta \implies T \vdash \alpha \text{ and } T \vdash \beta \implies T \vdash \delta \\
\mathfrak{N} \models \neg \delta & \implies \mathfrak{N} \models \neg \alpha \text{ or } \mathfrak{N} \models \neg \beta \implies T \vdash \neg \alpha \text{ or } T \vdash \neg \beta \implies T \vdash \neg \delta
\end{align*}
\]

where, for both parts, the central implication uses the induction hypothesis.

The second case is similar.
For the third case we have
\[ \mathfrak{A} \models \delta \implies \mathfrak{A} \models \neg \alpha \lor \mathfrak{A} \models \beta \implies T \models \neg \alpha \lor T \models \beta \implies T \models \delta \]
\[ \mathfrak{A} \models \neg \delta \implies \mathfrak{A} \models \alpha \quad \text{and} \quad \mathfrak{A} \models \neg \beta \implies T \models \alpha \quad \text{and} \quad T \models \neg \beta \implies T \models \neg \delta \]
where again the central implications use the induction hypothesis.

For the fourth case we have
\[ \mathfrak{A} \models \delta \implies \mathfrak{A} \models \neg \gamma \implies T \models \neg \gamma \implies T \models \delta \]
\[ \mathfrak{A} \models \neg \delta \implies \mathfrak{A} \models \gamma \implies T \models \gamma \implies T \models \neg \delta \]
where the central implications use the induction hypothesis. This is why we verify the two implications in tandem.

Another way to prove
\[ \mathfrak{A} \models \delta \implies T \models \delta \]
to rephrase \( \delta \) as a conjunction of disjunctions of literals, and then unravel this shape.

(b) We know that \( T \subseteq T^+ \subseteq Th(\mathfrak{N}) \). For the converse, consider any sentence \( \sigma \). We have
\[ T \models \sigma \iff \delta \]
for some quantifier-free sentence \( \delta \). Then
\[ \mathfrak{A} \models \sigma \implies \mathfrak{A} \models \delta \implies T \models \delta \implies T \models \sigma \]
using part (a), to give the required result. ■

2.14 Gödel’s theorem applied only to axiomatizations that capture at least a modicum of arithmetic. The theory \( Th(\mathbb{N}, S, 0) \) is too weak for this. It captures almost nothing of arithmetic. ■

B.4 Some other examples—not yet done
### C
#### Basic methods

**C.1 Some semantic relations**

3.2 Suppose $\mathfrak{A} \prec_r \mathfrak{B}$ and $\mathfrak{A} \models \sigma$ where $\sigma$ is a $\exists_{r+1}$-sentence. We have

$$\sigma = (\exists v) \phi(v)$$

for some $\forall_r$-formula $\phi(v)$, and there is some point $a$ of $\mathfrak{A}$ with $\mathfrak{A} \models \phi(a)$. But now $\mathfrak{A} \prec_r \mathfrak{B}$ gives $\mathfrak{B} \models \phi(a)$ so that $\mathfrak{B} \models \sigma$, as required. ■

**C.2 The diagram technique**

3.3 Let $a$ be an enumeration of the given structure. It suffices to show that

$$\Sigma \cup (Th(\mathfrak{A}, a) \cap \forall_1)$$

is consistent. This, of course, is a set of sentences in the enriched language.

By way of contradiction suppose this is not consistent. Then, by compactness, we have

$$\Sigma \vdash \neg \theta(a)$$

where $\theta(v)$ is a $\forall_1$-formula with $\mathfrak{A} \models \theta(a)$ for some point $a$ from $\mathfrak{A}$. This point $a$ does not occur in $\Sigma$, and so behaves like the batch $v$ of variables. Thus

$$\Sigma \vdash (\forall v) \neg \theta(v)$$

and hence

$$\mathfrak{A} \models (\forall v) \neg \theta(v)$$

by the given relationship between $\Sigma$ and $\mathfrak{A}$. The point $a$ ensures this is a contradiction. ■

3.4 Given such a pair of embeddings, for each $\forall_{n+1}$-sentence we have

$$\mathfrak{A} \models \sigma \implies \mathfrak{C} \models \sigma \implies \mathfrak{B} \models \sigma$$

using first $f$ then $g$.

Conversely, suppose that $\mathfrak{A} \equiv (\forall_{n+1}) \mathfrak{B}$ holds. Let $a$ be an enumeration of $\mathfrak{A}$ and let $b$ be an enumeration of $\mathfrak{B}$. We enrich the underlying language $L$ to both $L(a)$ and $L(b)$ and combine these to obtain $L(a, b)$. We arrange that the $a$-parameters and the $b$-parameters are disjoint. Look at

$$Th(\mathfrak{A}, a) \quad \Delta(\mathfrak{B}, b) = Th(\mathfrak{B}, b) \cap \forall_n$$

175
which are, respectively, a set of $L(a)$-sentences and a set of $L(b)$-sentences. These combine to give

$$Th(\mathfrak{A}, a) \cup \Delta(\mathfrak{B}, b)$$

a set of $L(a,b)$-sentences, and it suffices to show that this set is consistent.

By way of contradiction, suppose the set is not consistent. Then

$$Th(\mathfrak{A}, a) \vdash \neg \delta(b)$$

for some $\forall_n$-formular $\delta(w)$ and a point $b$ of $\mathfrak{B}$ with $\mathfrak{B} \models \delta(b)$. This gives first

$$Th(\mathfrak{A}, a) \vdash (\forall w)\neg \delta(w)$$

and then

$$\mathfrak{A} \models (\forall w)\neg \delta(w)$$

(since $a$ and $b$ are disjoint). But $(\forall w)\neg \delta(w)$ is a $\forall_{n+1}$-sentence (of the original language), and hence

$$\mathfrak{B} \models (\forall w)\neg \delta(w)$$

which leads to the contradiction. ■

3.5 One direction is immediate.

For the other direction suppose $\mathfrak{A} \equiv \mathfrak{B}$. Let $a, b$ be enumerations of these two structures and ensure that the corresponding families of constants are disjoint. It suffices to show that

$$Th(\mathfrak{A}, a) \cup Th(\mathfrak{B}, b)$$

is consistent.

If this set is not consistent then we have

$$\vdash \neg (\alpha(a) \land \beta(b))$$

for some pair $\alpha(u), \beta(v)$ of formulas where

$$\mathfrak{A} \models \alpha(a) \quad \mathfrak{B} \models \beta(b)$$

for points $a$ and $b$. The two batches $u, v$ are disjoint, so we have

$$\vdash (\exists u)\alpha(u) \rightarrow (\forall v)\neg \beta(v)$$

be some trivial manipulations. Since the point $a$ ensures that

$$\mathfrak{A} \models (\exists u)\alpha(u)$$

we have

$$\mathfrak{A} \models (\forall v)\neg \beta(v)$$

so that

$$\mathfrak{B} \models (\forall v)\neg \beta(v)$$

which is contradicted by the point $b$. ■
C.3 Restricted axiomatization

3.6 (a) This is a generalization of Lemma 3.16. One direction is immediate. Conversely, suppose

\[ \mathfrak{A} \models T \cap \forall_{n_1} \]

and let \( a \) an enumeration of \( \mathfrak{A} \). It suffices to show that

\[ T \cup (\text{Th}(\mathfrak{A}, a) \cap \forall_n) \]

is consistent.

If this is not consistent then

\[ T \vdash \neg \alpha(a) \]

for some \( \forall_n \)-formula \( \alpha(v) \) where

\[ \mathfrak{A} \models \alpha(a) \]

for some matching point \( a \). The fist of these gives

\[ T \vdash (\forall v) \neg \alpha(v) \]

and hence

\[ \mathfrak{A} \models (\forall v) \neg \alpha(v) \]

which is not so.

(b) This is a generalization of Theorem 3.17. One direction is immediate. Conversely, suppose the theory \( T \) has the indicated property. We show that

\[ \mathfrak{A} \models T \cap \forall_{n+1} \implies \mathfrak{A} \models T \]

for each structure \( \mathfrak{A} \).

For each structure

\[ \mathfrak{A} \models T \cap \forall_{n+1} \]

by part (a) we have some structure \( \mathfrak{B} \) with

\[ \mathfrak{A} \preceq_n \mathfrak{B} \models T \]

and then

\[ \mathfrak{A} \models T \]

by the given property of \( T \).

(c) This is a generalization of Theorem 3.18. The implication \((i) \Rightarrow (ii)\) is straight forward.

For the other implication we assume \((ii)\) and let \( \Sigma \) be the set of \( \forall_{n+1} \)-sentences \( \sigma \) such that

\[ T \vdash \lambda \rightarrow \sigma \]

holds. We also let \( T' \) be the theory axiomatized by \( T \cup \{ \lambda \} \). Thus \( \Sigma = T \cap \forall_{n+1} \).

Consider any model \( \mathfrak{A} \models \Sigma \). By part (a) we have

\[ \mathfrak{A} \preceq_n \mathfrak{B} \models T' \]
C. Basic methods

for some structure $\mathfrak{B}$. We have $\mathfrak{B} \models \lambda$, and hence (ii) gives $\mathfrak{A} \models \rho$. This shows that

$$\Sigma \vdash \rho$$

and hence a use of compactness gives the required result.

3.7 This proof is very like that of Solution 3.4, and generalizes the proof of Lemma 3.19. Of the required implications is immediate. For the other suppose $\mathfrak{A} \prec_{n+1} \mathfrak{B}$. Let $a$ be an enumeration of $\mathfrak{A}$, and let $b$ be an enumeration of $\mathfrak{B}$. It suffices to show that

$$Th(\mathfrak{A}, a) \cup (Th(\mathfrak{B}, a, b) \cap \forall_n)$$

is consistent. The left hand component is a set of $L(a)$ sentences. The right hand component is a set of $L(a, b)$-sentences, and will contains some sentences ($a \equiv b$) where $a$ and $b$ are parameters that name the same element of $\mathfrak{B}$.

By way of contradiction suppose the displayed set is not consistent. Then

$$Th(\mathfrak{A}, a) \vdash \neg \delta(b, a)$$

for some $\forall_n$-formula $\delta(v, u)$ of the parent language, an points $a$ from $a$ and $b$ from $b$. This gives $\mathfrak{A} \models (\forall v) \neg \delta(v, a)$ and then $\mathfrak{B} \models (\forall v) \neg \delta(v, a)$ (since $\mathfrak{A} \prec_{n+1} \mathfrak{B}$), which leads to the contradiction.

3.8 This is a kind of generalization of Theorem 3.18. We set up the proof in a slightly different fashion.

$(i) \Rightarrow (ii)$. Assuming $(i)$ consider any pair $\mathfrak{A} \subseteq \mathfrak{B}$ of models of $T$ and let $a$ be a point of $\mathfrak{A}$ (and hence also of $\mathfrak{B}$). We have

$$\mathfrak{A} \models \phi(a) \rightarrow \theta(a) \quad \mathfrak{B} \models \theta(a) \rightarrow \psi(a)$$

so that

$$\mathfrak{A} \models \phi(a) \implies \mathfrak{A} \models \theta(a) \implies \mathfrak{B} \models \theta(a) \implies \mathfrak{B} \models \psi(a)$$

since $\theta$ is an $\exists_1$-formula.

$(ii) \Rightarrow (i)$. Assuming $(ii)$ let $v$ be the batch of variable $\partial\phi \cup \partial\psi$. Let $\Pi(T, \neg \psi)$ be the set of all $\forall_1$-formulas $\alpha(v)$ such that

$$T \vdash \neg \psi \rightarrow \alpha$$

and only the batch $v$ occurs in $\alpha$. This set $\Pi(T, \neg \psi)$ is a $\forall_1$-type over $T$. It suffices to show that

$$(?) \quad T \cup \Pi(T, \neg \psi) \vdash \neg \phi$$

for then we have

$$T \vdash \alpha \rightarrow \neg \phi$$

for some $\alpha \in \Pi(T, \neg \psi)$, and hence $\theta = \neg \alpha$ gives us $(i)$.

By way of an outer contradiction suppose that $(?)$ does not hold. Then

$$T \cup \Pi(T, \neg \psi) \cup \{\phi\}$$

is consistent to give us

$$\mathfrak{A} \models T \quad \mathfrak{A} \models \Pi(a) \quad \mathfrak{A} \models \phi(a)$$
for some structure $\mathfrak{A}$ and point $a$ matching the batch $v$. Let $a$ be an enumeration of $\mathfrak{A}$ and consider the set of sentences

$$(!) \quad T \cup \text{Diag}(\mathfrak{A}, a, a) \cup \{\neg \psi(a)\}$$

in the enriched language. (It is true that $a$ will occur as a part of $a$ but it is more convenient to also have a distinguished version.) We show that ($!$) is consistent.

By way of an inner contradiction\(^1\) suppose that ($!$) is consistent. Then we get a structure $\mathfrak{B}$ with

$$\mathfrak{B} \models T \quad \mathfrak{A} \subseteq \mathfrak{B} \quad \mathfrak{B} \models \neg \psi(a)$$

and hence the hypothesis (ii) gives

$$\mathfrak{A} \models \neg \phi(a)$$

which is the inner contradiction.

Since ($!$) is not consistent a use of compactness gives

$$T \vdash \delta(a, x) \rightarrow \psi(a)$$

for some QF formula $\delta(v, w)$ with

$$\mathfrak{A} \models \delta(a, x)$$

for some point $x$ of $\mathfrak{A}$ matching the batch $w$. By transfering to variables we have

$$T \vdash \delta(v, w) \rightarrow \psi(w)$$

so that

$$T \vdash (\exists w) \delta(v, w) \rightarrow \psi(w)$$

to give

$$\neg (\exists w) \delta(v, w) \in \Pi(T, \neg \psi)$$

which leads to the outer contradiction by our choice of $\mathfrak{A}$. ■

### C.4 Directed families of structures

3.9 We verify this by induction on $n$.

The base case, $n = 0$, is trivial.

For the induction step, $n \mapsto n + 1$, suppose

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec_{n+1} \mathfrak{B}$$

for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{A}$. Consider any $\mathfrak{A} \in \mathcal{A}$, and suppose

$$\mathfrak{A} \models (\forall v) \phi(v, a)$$

where $\phi(v, u)$ is a $\exists_n$-formula, and $a$ is a point from $\mathfrak{A}$. To show

$$\mathfrak{A} \models (\forall v) \phi(v, a)$$

\(^1\)This kind of proof really upsets some people. There is only one answer to that: Good
consider any point \( b \) from \( U \) which matches \( v \). There is some \( B \in A \) with \( A \subseteq B \) and such that \( b \) come from \( B \). But now \( A \prec_{n+1} B \), so that
\[
B \models (\forall v)\phi(v, a)
\]
and hence
\[
B \models \phi(b, a)
\]
holds.

The induction hypothesis gives \( B \prec_n U \), and hence
\[
U \models \phi(b, a)
\]
as required. ■

3.10 Using induction on \( n \) we show that
\[
[n] \quad A \subseteq B \implies A \prec_n B
\]
holds for all \( A, B \in A \) (quantified inside the induction hypothesis).

The base case, \([0]\), is trivial.

For the induction step, \([n] \implies [n + 1]\), consider \( A, B \in A \) with \( A \subseteq B \). The condition on \( A \) gives
\[
A \prec C \quad B \subseteq C
\]
for some \( C \in A \). The induction hypothesis applied to \( B, C \) gives \( B \prec_n C \), and hence \( A \prec_{n+1} B \), as required. ■

3.11 Consider any member \( a \) of the carrier of \( \bigcup A \). Then \( a \) belongs to some \( A \in A \). But now \( A \subseteq B \) for some \( B \in B \). Thus \( a \) belongs to the carrier of \( \bigcup B \). This with a similar argument shows that the carriers of \( \bigcup A \) and \( \bigcup B \) are the same.

The same kind of reasoning sows that they carry the same attributes. ■

3.12 It suffices to extract a bit more information out of the proof of Lemma 3.25. That proof provides models \( B, C, D \) of \( T \) with
\[
U \prec_1 B \prec D \quad U \subseteq C \subseteq E \quad U = B \cap E
\]
where \( U \subseteq B \) holds by the choice of \( B \). In particular, we have \( U \prec_1 D \), and hence Exercise 3.7 gives a structure \( E \) with \( D \subseteq E \) and \( U \prec E \), as required. ■

C.5 The up and down techniques

3.13 Under the given conditions we show
\[
A \prec_n B
\]
by induction on \( n \). The base case, \( n = 0 \), is trivial. For the induction step, \( n \mapsto n + 1 \), suppose \( B \models (\exists v)\phi(v, a) \) where \( \phi(v, u) \) is a \( \forall_n \)-formula, and \( a \) is a point from \( A \). We do not assume that \( v \) is a single variable, since iterated use of the given condition gives \( B \models \phi(b, a) \) for some point \( b \) also from \( A \). The induction hypothesis, \( A \prec_n B \), now gives \( A \models \phi(b, a) \), so that \( A \models (\exists v)\phi(v, a) \), as required. ■
Model complete and submodel complete theories

D.1 Model complete theories

4.1 Suppose first that the theory $T$ is model complete and consider any pair $\mathfrak{A} \subseteq \mathfrak{B}$ of models. We know that the enriched theory $T[\mathfrak{A}]$ is complete. In more detail the theory axiomatized by

$$T(\mathfrak{A}) = T \cup \text{Diag}(\mathfrak{A}, a)$$

is complete. Here $a$ is an enumeration of $\mathfrak{A}$. (We might ask which enumeration is this. It doesn’t matter which one. If what follows works for one, then it works for all.) This completeness ensures that any sentence of the enriched language which is consistent with the set $T(\mathfrak{A})$ holds in all models of the set. Since $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{B}$ is a model of $T$, we see that $(\mathfrak{B}, a)$ is a model of $T(\mathfrak{A})$.

Now suppose

$$\mathfrak{A} \models \phi(a)$$

for some formula $\phi(v)$ and point $v$ of $\mathfrak{A}$ that matches the batch $v$ of variables. We view $\phi(a)$ as a sentence of the enriched language. Since this sentence holds in $(\mathfrak{A}, a)$, it is consistent with $T(\mathfrak{A})$, and hence holds in $(\mathfrak{B}, a)$. In other words we have

$$\mathfrak{B} \models \phi(a)$$

in terms of the original language.

This shows that $\mathfrak{A} \prec \mathfrak{B}$, as required.

Conversely, suppose we know that

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \prec \mathfrak{B}$$

for all models $\mathfrak{A}, \mathfrak{B}$ of $T$. Consider any particular model $\mathfrak{A}$ and look at the set $T(\mathfrak{A})$, as above. We must show that this set axiomatizes a complete theory. In other words we must show that any sentence (in the enriched language) which is consistent with $T(\mathfrak{A})$ automatically holds in all models of $T(\mathfrak{A})$.

Consider any such sentence. We view that as $\phi(a)$ for some formula $\phi(v)$ of the parent language and some point $a$ of $\mathfrak{A}$ which matches the batch $v$. Since $\phi(a)$ is consistent with $T(\mathfrak{A})$, it holds in some model of $T(\mathfrak{A})$. Thus we have

$$\mathfrak{A} \subseteq \mathfrak{B} \models \phi(a)$$

for some model $\mathfrak{B}$ of $T$.

By the supposition we have $\mathfrak{A} \prec \mathfrak{B}$ and hence

$$\mathfrak{A} \models \phi(a)$$
(since \(a\) is a point of \(A\)).

Now consider any model \((C, a)\) of \(T(A)\). As before we have \(A \subseteq C\), so that \(A < C\), and hence
\[
\mathcal{C} \models \phi(a)
\]
or equivalently, the sentence \(\phi(a)\) holds in the arbitrary model \((C, a)\) of \(T(A)\).

4.2 For the first part we verify each \([n]\) by induction on \(n\). But to do that we must correct the statement of each. We can not fix the models \(A, B\) throughout, they must be allowed to vary in the argument. Thus \([n]\) should have the form

For all models \(A, B\) we have \(\ldots\)

with the quantification inside.

With this rephrasing, we assume \([0]\) and verify \([n]\) by induction on \(n\).

The base case, \([0]\), is given. For the induction step, \(n \mapsto n + 1\), suppose we have

\[
A <_{n+1} B
\]

for a pair \(A, B\) of models of \(T\). By Exercise 3.7 we have

\[
\begin{align*}
A & \subseteq B \\
A & <_{n} B \\
& \hspace{1cm} <_{n+1} C
\end{align*}
\]

for some structure \(C\). In particular, \(C\) is a model of \(T\), and so the right hand inclusion and the induction hypothesis \([n]\) gives \(B <_{n+1} C\). Another use of Exercise 3.7 (in the other direction) now gives \(A <_{n+2} B\), as required.

For the second part suppose the theory has \([0]\), and consider any pair \(A \subseteq B\) of models. By the first part we have \(A <_{n} B\) for each \(n\), and hence \(A < B\), as required. ■

4.3 (a) Let \(T\) be a theory with \(EQ\) and consider any pair \(A \subseteq B\) of models of \(T\). We show that \(A \prec B\).

Suppose

\[
A \models \phi(a)
\]

for some formula \(\phi(v)\) and object \(a\) of \(A\) matching the batch \(v\) of variables. Since \(T\) has \(EQ\) we have

\[
T \vdash (\forall v)[\phi(v) \leftrightarrow \delta(v)]
\]

for some quantifier-free formula \(\delta(v)\). But now we have

\[
A \models \delta(a)
\]
to give

\[
B \models \delta(a)
\]

and hence

\[
B \models \phi(a)
\]

for the required result.

(b) There are two slightly different proofs of this that should be known.
The first proof uses the characterization of Theorem 3.24. Let $T$ be a model complete theory, and consider any ascending $\omega$-chain

$$A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots$$

of models of $T$, and let $\mathcal{B}$ be the union of this chain. We require $\mathcal{B} \models T$. But the model completeness gives

$$A_0 \prec A_1 \prec A_2 \prec \cdots$$

and hence $A_0 \prec \mathcal{B}$ by Lemma 3.23, to give the required result.

For the second proof consider any structure $A \prec \mathcal{B}$ where $\mathcal{B} \models T$. We show

$$[n] A \prec_{2n+1} \mathcal{B}$$

by induction on $n$. Thus gives $A \prec \mathcal{B}$, and hence $A \models T$.

The base case, $[0]$, is trivial.

For the induction step, $n \mapsto n + 1$, we have

$$A \prec_{2n+1} \mathcal{B}$$

by the induction hypothesis. A use of the fundamental construction of Exercise 3.7 gives

$$A \prec_{2n+1} \mathcal{B} \prec_{2n} \mathcal{C}$$

for some structure $\mathcal{C}$. Since $\mathcal{B} \models T$ we have $A \models T \cap \forall_{2n+2}$ and hence $\mathcal{C} \models T \cap \forall_{2n+2}$ so that

$$\mathcal{C} \prec_{2n+1} \mathcal{D}$$

for some model $\mathcal{D}$ of $T$. Thus we have

$$A \prec_{2n+1} \mathcal{B} \prec_{2n} \mathcal{C} \prec_{2n+1} \mathcal{D}$$

where the upper elementary embedding holds by the model completeness of $T$. But now moving from right to left we have

$$\mathcal{C} \prec_{2n+1} \mathcal{D}$$

to give

$$\mathcal{B} \prec_{2n+2} \mathcal{C}$$

and hence

$$A \prec_{2n+3} \mathcal{B}$$

as required. \[\blacksquare\]
D.2 Two structural properties

4.4 We must show that if the theory $T$ has $JEP$ then

$$T \vdash \alpha \lor \beta \implies T \vdash \alpha \text{ or } T \vdash \beta$$

for each pair $\alpha, \beta$ of $\forall_1$-sentences. In fact we prove the contrapositive.

Suppose

$$T \not\vdash \alpha \quad T \not\vdash \beta$$

that is both

$$T \cup \{\neg \alpha\} \quad T \cup \{\neg \beta\}$$

are consistent to give

$$\mathfrak{A} \models \neg \alpha \quad \mathfrak{B} \models \neg \beta$$

for a pair $\mathfrak{A}, \mathfrak{B}$ of models of $T$. Since $T$ has $JEP$ these can be embedded into some model $\mathfrak{C}$ of $T$ and then

$$\mathfrak{C} \models \neg \alpha \land \neg \beta$$

since both these sentences are $\exists_1$. This gives

$$T \not\vdash (\alpha \lor \beta)$$

as required.

4.5 This has two different proofs, both of which are instructive.

For the first proof observe that the characterization of Theorem 4.6(ii) depends only $T \cap \forall_2$.

For the second proof we show that if $T$ has $AP$ then so does $T \cap \forall_2$. (The other required implication is immediate.)

Consider any wedge of embeddings between models of $T \cap \forall_2$, as in (1) of Table D.1. By enlarging $\mathfrak{B}, \mathfrak{C}$ we may suppose that both are models of $T$. A similar trick gives us a diagram as in (2) of Table D.1 where $\mathfrak{A}', \mathfrak{B}, \mathfrak{C} \in \mathcal{M}(T)$ and $l$ is a $\prec_1$-embedding. (This is because $\mathfrak{A} \models T \cap \forall_2$.) Two uses of Lemma 4.7 gives us a commuting diagram of embeddings as in (3) of Table D.1 where $m$ and $n$ are $\prec$-embeddings. In particular, $\mathfrak{A}', \mathfrak{B}', \mathfrak{C}' \in \mathcal{M}(T)$. But $T$ has $AP$, so the right hand wedge can be closed to give a diagram as in (4) of Table D.1 where $\mathfrak{D}' \models T$.

4.6 This is an immediate consequence of the more general result of Lemma 4.7.

4.7 Suppose $T \cap \forall_1$ has $AP$, and consider formulas $\theta(v) \in \exists_1$ and $\phi(v) \in \forall_1$ (in the indicated variables) such that

$$T \vdash \theta(v) \rightarrow \phi(v)$$

holds. Let $\Delta(v)$ be the set of quantifier-free formulas $\delta(v)$ such that

$$T \vdash \theta(v) \rightarrow \delta(v)$$

holds. If suffices to show that

$$T \cup \Delta(v) \vdash \phi(v)$$

holds.
Consider any model \((\mathcal{B}, a)\) of the hypothesis set. Thus
\[
\mathcal{B} \models T \quad \mathcal{B} \models \Delta(a)
\]
for some point \(a\) of \(\mathcal{B}\). We must show that \(\mathcal{B} \models \phi(a)\).

Let \(\mathfrak{A}\) be the substructure of \(\mathcal{B}\) generated by \(a\). Thus \(\mathfrak{A} \models T \cap \forall_1\). A simple argument shows that
\[
T \cup \text{Diag}(\mathfrak{A}, a) \cup \{\theta(a)\}
\]
is consistent. Thus we have a wedge of embeddings as in (1) of Table D.1 where \(\mathcal{C} \models T\) and \(\mathcal{C} \models \theta(ga)\). Here \(f\) is the insertion of \(\mathfrak{A}\) into \(\mathcal{B}\). Since \(T \cap \forall_1\) has \(AP\), this closes to a commuting square

![Diagram](image)

where \(\mathfrak{D} \models T \cap \forall_1\). In fact, by a suitable enlargement we can arrange that \(\mathfrak{D} \models T\).

We have
\[
\mathcal{C} \models \theta(ga) \quad T \vdash \theta \rightarrow \phi
\]
and \(\theta \in \exists_1\), so that
\[
\mathfrak{D} \models \theta((k \circ g)a) \quad T \vdash \theta \rightarrow \phi
\]
to give
\[
\mathfrak{D} \models \theta((h \circ f)a)
\]
since \(h \circ f = k \circ g\). But \(\phi \in \forall_1\), so that
\[
\mathcal{B} \models \phi(fa)
\]
and hence
\[
\mathcal{B} \models \phi(a)
\]
since $f$ is an insertion. This is the required result.

4.8 Suppose first that $\mathfrak{A}$ is an amalgamation base for $T$, and suppose

$$T \vdash \psi \lor \phi$$

where $\psi(v)$ and $\phi(v)$ are $\forall_1$-formulas. By way of contradiction, suppose $\mathfrak{A}$ realizes the associated types at some point $a$. Let $a$ be an enumeration of the whole of $\mathfrak{A}$. We check that both

$$T \cup \text{Diag}(\mathfrak{A}, a) \cup \{\neg \psi(a)\} \quad T \cup \text{Diag}(\mathfrak{A}, a) \cup \{\neg \phi(a)\}$$

are consistent, and hence we obtain a wedge of embeddings as in (1) of Table D.1 where

$$\mathfrak{B} \models \neg \psi(a) \quad \mathfrak{C} \models \neg \phi(a)$$

and with $\mathfrak{B}, \mathfrak{C} \in \mathcal{M}(T)$. Since $\mathfrak{A}$ is an amalgamation base, we may close this wedge, and then argue as in the proof of Theorem 4.6, (i)$\Rightarrow$(ii).

Conversely, suppose $\mathfrak{A} \in \mathcal{S}(T)$ has the omitting property, and consider a wedge of embeddings as in (1) of Table D.1 where $\mathfrak{B}, \mathfrak{C} \in \mathcal{M}(T)$. As in the proof of Theorem 4.6, (ii)$\Rightarrow$(i) we may suppose

$$\mathfrak{A} \subseteq \mathfrak{B} \quad \mathfrak{A} \subseteq \mathfrak{C} \quad \mathfrak{B} \cap \mathfrak{C} = \mathfrak{A}$$

hold. Let $a, b, c$ be enumerations of $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$ respectively. It suffices to show that

$$T \cup \text{Diag}(\mathfrak{B}, b, a) \cup \text{Diag}(\mathfrak{C}, c, a)$$

is consistent.

By way of contradiction, suppose this set is not consistent. Then we find $\forall_1$-formulas $\psi(u), \phi(u)$ such that

$$T \vdash \psi \lor \phi \quad \mathfrak{B} \models \neg \psi(a) \quad \mathfrak{A} \models \neg \phi(a)$$

for some point $a$ of $\mathfrak{A}$. This point $a$ can not realize

$$\neg \Sigma(T, \psi) \cup \neg \Sigma(T, \phi)$$

in $\mathfrak{A}$. Suppose it does not realize $\neg \Sigma(T, \psi)$, say. Then we have

$$\mathfrak{A} \models \lambda(a) \quad T \vdash \lambda \to \psi$$

for some $\exists_1$-formula $\lambda(u)$, and hence $\mathfrak{B} \models \lambda(a)$, which leads to the contradiction.

4.9 (a) This is almost immediate.

(b) This is prove in the same way as the argument of Solution 4.5.

(c) This is an immediate consequence of the characterization given in Exercise 4.8.

D.3 Submodel complete theories

4.10 Let $T$ be a theory which is both model complete and $\forall_1$-axiomatizable. By Exercise 4.6 (or Lemma 4.7) we see that $T$ has $\text{AP}$. But then $T \cap \forall_1$ has $\text{AP}$ (since this set axiomatizes $T$). Thus $T$ has $\text{EQ}$ be Theorem 4.10.
E
The back and forth technique

E.1 The technique

5.1 The proof that each of $\cong, \equiv, \equiv_n$ is an equivalence relation is more or less immediate. The proof for $\equiv_p$ is not so straightforward.

Consider any structure $\mathfrak{A}$. There is at least one isomorphism

$$f : \mathfrak{A} \longrightarrow \mathfrak{A}$$

for we can take the identity function. The singleton $\{f\}$ is a b&f-system, to show that $\equiv_p$ is reflexive.

Almost trivially, $\equiv_p$ is symmetric.

The transitivity of $\equiv_p$ is not so immediate.

Before we show that we make a couple of observations about the nature of b&f systems.

Suppose we have

$$\mathfrak{A} \equiv_p \mathfrak{B}$$

for a pair of structures $\mathfrak{A}, \mathfrak{B}$. This is witnessed by some b&f system $P$ which is a family of certain bijections

$$f : U \longrightarrow V$$

between subsets $U$ of $\mathfrak{A}$ and $V$ of $\mathfrak{B}$.

Let $P_1$ be the set of all restrictions

$$f_{\mid F} : F \longrightarrow f[F]$$

of such $f \in P$ to finite sets $F$ (where $F \subseteq U$ for this $f$). A short cogitation shows that $P_1$ is also a b&f-system between $\mathfrak{A}, \mathfrak{B}$. In other words, if

$$\mathfrak{A} \equiv_p \mathfrak{B}$$

then this relationship can be witness by a system of finite partial isomorphisms and where $P$ is closed under taking restrictions.

With this we can quickly show that $\equiv_p$ is transitive.

Suppose we have

$$\mathfrak{A} \equiv_p \mathfrak{B} \equiv_p \mathfrak{C}$$

witnessed by b&f systems

$$P \quad Q$$

respectively. As explained above, we may assume that each of $P, Q$ consists of finite partial isomorphisms and is closed under taking restrictions.

Consider any finite subset $U$ of $\mathfrak{A}$. By starting from any member of $P$, repeatedly going forth to bring $U$ into the fold, and then restricting to $U$, we see there is at least one partial isomorphism

$$f : U \longrightarrow V$$
in $P$. With this finite subset $V$ of $\mathfrak{B}$ there is at least one partial isomorphisms

$$g : V \rightarrow W$$

in $Q$. The composite

$$g \circ f : U \rightarrow W$$

is a finite partial isomorphism.

Let

$$Q \circ P$$

be the family of all the partial isomorphisms $g \circ f$ obtained in this way. It doesn’t take long to see that $Q \circ P$ is a b&f system, and hence

$$\mathfrak{A} \equiv_p \mathfrak{C}$$

as required. 

5.2 This can be obtained by modifying the proof of Theorem 5.5. However, there is no need to do that.

Let $P$ be a b&f-system between $\mathfrak{A}, \mathfrak{B}$. For an arbitrary $(a, b) \in P$, consider the pair

$$(\mathfrak{A}, a) (\mathfrak{B}, b)$$

structures. A few moments thought shows that $P$ is a b&f-system between these enriched structures. Thus by invoking Theorem 5.5 we have

$$(\mathfrak{A}, a) \equiv (\mathfrak{B}, b)$$

as required.

5.3 The $\aleph_0$-categoricity is an immediate consequence of Theorem 5.6.

For the completeness consider any pair $\mathfrak{A}', \mathfrak{B}'$ of models. We show $\mathfrak{A}' \equiv \mathfrak{B}'$.

Each of these structures is infinite. By the downwards LS-theorem we have

$$\mathfrak{A} \prec \mathfrak{A}' \quad \mathfrak{B} \prec \mathfrak{B}'$$

for two countable structures $\mathfrak{A}, \mathfrak{B}$. Each of these is a model of the theory, so that

$$\mathfrak{A} \equiv_p \mathfrak{B}$$

to give

$$\mathfrak{A} \equiv \mathfrak{B}$$

and hence

$$\mathfrak{A}' \equiv \mathfrak{A} \equiv \mathfrak{B} \equiv \mathfrak{B}'$$

to give

$$\mathfrak{A}' \equiv \mathfrak{B}'$$

as required.
E.2 Linear orders

5.5 The cartesian product $B \times A$ consists of ordered pairs $(b, a)$ where $a \in A$ and $b \in B$. These are ordered by

$$(b_1, a_1) \leq (b_2, a_2) \iff (a_1 < a_2) \text{ or } ((a_1 = a_2) \text{ and } (b_1 \leq b_2))$$

thus we look first at the $A$-component and if that doesn’t separate the pair then we look at the $B$-component.

5.6 (a) The associative law is more or less trivial.

The two linear orders

$$\mathbb{N} + \mathbb{Q}, \mathbb{Q} + \mathbb{N}$$

are nothing like each other to show that addition is not commutative.

(b) To show that multiplication is associative we unravel the comparisons

$$(c_1, (b_1, a_1)) \leq (c_2, (b_2, a_2)) \quad ((c_1, b_1), a_1) \leq ((c_2, b_2), a_2)$$

using linear ordered sets $A, B, C$.

For the left hand comparison the first unravel gives

$$(b_1, a_1) < (b_2, a_2) \text{ or } ((b_1, a_1) = (b_2, a_2) \text{ and } (c_1 \leq c_2))$$

and this gives

$$(a_1 < a_2) \text{ or } ((a_1 = a_2) \text{ and } (b_1 < b_2)) \text{ or } ((a_1 = a_2) \text{ and } (b_1 = b_2) \text{ and } (c_1 \leq c_2))$$

which we leave as it is.

For the right hand comparison the first unravel gives

$$(a_1 < a_2) \text{ or } ((a_1 = a_2) \text{ and } ((c_1, b_1) \leq (c_2, b_2)))$$

and this gives

$$(a_1 < a_2) \text{ or } ((a_1 = a_2) \text{ and } ((b_1 = b_2) \text{ and } (c_1 \leq c_2)))$$

which we compare with the previous expansion. A few boolean manipulations shows that the two expansions are equivalent.

Let $2$ be the 2-element linearly ordered set. By drawing a couple of pictures we see that

$$\mathbb{Z} \times 2, 2 \times \mathbb{Z}$$

are

$$\mathbb{Z} + \mathbb{Z}, \mathbb{Z}$$

which are not the same, and so multiplication is not commutative.

(c) It doesn’t take too long to show

$$\gamma \times (\beta + \alpha) = \gamma \times \beta + \gamma \times \alpha$$

for arbitrary order types $\alpha, \beta, \gamma$. 
The other distributive law
\[(\gamma + \beta) \times \alpha = \gamma \times \alpha + \beta \times \alpha\]
does not hold in general. To see this let 1 be the 1-element set. Thus
\[1 + 1 = 2\]
so that
\[(1 + 1) \times \mathbb{Z} = 2 \times \mathbb{Z} = \mathbb{Z} \quad 1 \times \mathbb{Z} + 1 \times \mathbb{Z} = \mathbb{Z} + \mathbb{Z}\]
to give a counter-example.

E.3 Graphs

5.7 For each pair \(l, r \in \mathbb{N}\) let \(\rho(l, r)\) be the \(\forall_2\)-sentence
\[(\forall u_0, \ldots, u_l, w_0, \ldots, w_r)(\exists v) [\bigwedge \{ u_i \sim v \mid 0 \leq i \leq l \} \land \bigwedge \{ v \not\sim w_j \mid 0 \leq j \leq r \}]\]
in the language of graphs. The set of all such sentence axiomatizes the notion of randomness (for graphs).

5.8 We show
\[V_r \subseteq V_{r+1}\]
by induction on \(r\). The base case, \(r = 0\), is trivial (since \(\emptyset\) is a subset of every set). For the induction step, \(r \mapsto r + 1\), we assume
\[V_r \subseteq V_{r+1}\]
and consider an arbitrary \(x \in V_{r+1}\). Then
\[x \subseteq V_r \subseteq V_{r+1}\]
to give
\[x \in \mathcal{P}V_{r+1} = V_{r+2}\]
and hence
\[V_{r+1} \subseteq V_{r+2}\]
as required.

For \(V_0, V_1, V_2\) we have
\[
\begin{array}{ccc}
V_0 & V_1 & V_2 \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\end{array}
\]
where for \(V_1, V_2\) we have only drawn the new trees. For \(V_3\) let’s draw all the trees including the once from \(V_2\). These are displayed in Table E.1.
Table E.1: Some small sets as trees

After that the size explodes.

5.9 There is no finite set $x$ with $x \in x$. ■

5.10 (a) The relation is trivially symmetric, and since no member $x$ of the ZF-hierarchy satisfies $x \in x$, the relation is irrelexive.

(b) We take $V = V_\lambda$ for some appropriate limit ordinal (to be determined). We require $V$ to satisfy

$$z \in V \implies |z| < \kappa \quad \{ z \subseteq V \mid |z| < \kappa \} \implies z \in V$$

and this can be arranged by taking $\lambda$ suitably large.

Consider any disjoint pair $L, R$ subsets of $V$ with $|L|, |R| < \kappa$. By the previous remark we have $L, R \in V$. We consider

$$a = L \cup b$$

for some suitable $b \subseteq V$ with $|b| < \kappa$. Again by the previous remark we have $a \in V$. Since

$$x \in L \implies x \in a \implies x \sim a$$
E. The back and forth technique

we immediately satisfy one of the requirements. We require that
\[ y \notin a \text{ and } a \notin y \]
for each \( y \in R \).

Since \( L \) and \( R \) are disjoint a choice of \( b \) with
\[ R \cap b = \emptyset \]
will ensure the right hand requirement. Thus any \( b \) with
\[ b \subseteq V - R \quad |b| < \kappa \]
will satisfy this requirement. Since \( |R| < \kappa \leq |V| \), we have
\[ |V - R| = |V| \]
this gives us plenty of choice for \( b \).

Suppose that \( a \in y \in R \) for some set \( y \). Let \( A \) be the set of all such \( a \). Since \( |R| < \kappa \) and each \( y \in R \) satifies \( |y| < \kappa \) we have \( |A| < \kappa \). This bars no strictly less that \( \kappa \) sets \( b \). But Since \( V - R \) has many more subsets, there is at least one choice for \( b \). \( \blacksquare \)

5.11 By way of contradiction supposed that none of the subgraphs \( U_1 \ldots, U_m \) is random. For each \( 1 \leq i \leq m \) let
\[ L_i, R_r \subseteq B_i \quad L_1 \cap R_i = \emptyset \]
witness this non-randomness. Thus
\[ \neg(\exists b \in U_i)[L_i \dashv \vdash b \dashv \vdash R_i] \]
using the obvious abbreviations.

Let
\[ L = L_1 \cup \cdots \cup L_m \quad R = R_1 \cup \cdots \cup R_m \]
to obtain two subsets of \( V \). Observe that
\[ L_i \cap R_j = \emptyset \]
for each pair of indexes \( 1 \leq i, j \leq m \). When \( i \neq j \) we have
\[ L_i \cap R_j \subseteq U_i \cap U_j = \emptyset \]
and
\[ L_1 \cap R_i = \emptyset \]
by the choice of these sets. This gives
\[ L \cap R = \emptyset \]
by a standard boolean manipulation.

Since \( V \) is random we have
\[ L \dashv \vdash a \dashv \vdash R \]
for some \( a \in V \). This node must live in some \( U_i \), and hence
\[ L_i \dashv \vdash a \dashv \vdash R_i \]
which is the contradiction. \( \blacksquare \)
E.4 Equivalence relations

5.12 We have to furnish the set

$$D = (B - A) \cup A \cup (C - A)$$

with an equivalence relation to produce a superstructure $D$ of both $B$ and $C$. To do that we consider what the block of $D$ should look like.

Consider any $a \in A$. This gives us three blocks in which $a$ lives, namely

$$[a]_B \quad [a]_A \quad [a]_C$$

the block in the structure

$$B \quad A \quad C$$

respectively. Of course, we have

$$A \cap [a]_B = [a]_A = B \cap [a]_C$$

since $A$ is a substructure of $B$ and $C$.

If

$$[a]_B = [a]_A = [a]_C$$

then we leave this set as a block of $D$.

If $[a]_B = [a]_A$ but $[a]_C$ meets $C - A$, then we leave $[a]_C$ as the block $[a]_D$ of $D$.

Similarly, if $[a]_C = [a]_A$ but $[a]_B$ meets $B - A$, then we leave $[a]_B$ as the block $[a]_D$ of $D$.

What do we do if both

$$[a]_B - A \quad [a]_C - A$$

are non-empty? We take

$$([a]_B - A) \cup ([a]_C - A)$$

as the block $[a]_D$ of $D$.

Consider any $b \in B$. We may suppose that $[b]_B$ does not meet $A$ (for otherwise we have already dealt with it). For this case we leave $[b]_B$ as the block $[b]_D$ of $D$.

Consider any $c \in C$. We may suppose that $[c]_C$ does not meet $A$ (for otherwise we have already dealt with it). For this case we leave $[c]_C$ as the block $[c]_D$ of $D$. ■

E.5 Miniature arithmetic

5.13 This is easier than it looks. We use the fact the model $\mathbb{N}$ of $T(\leq, S, 0)$ is uniquely embeddable in every submodel.

Consider a pair $L, R$ of submodels. Using $\mathbb{N}$ we have a wedge

$$\begin{tikzcd}
L \ar[dr] & R \\
\mathbb{N} \ar[ur]
\end{tikzcd}$$

and then a use of $AP$ embeds $L$ and $R$ into a model (for which the resulting square commutes). ■
F

Companion theories an existentially closed structures

F.1 Model companions

6.1 (a) Suppose $T^*$ is a model completion of $T$. Consider any model $A \models T^*$. Then $A \models T$, so that $T^*[A]$ is complete, be the definition of model comletion. This shows that $T^*$ is model complete.

Consider any wedge

of models of $T$. By enlargement we may suppose $B, C$ are models of $T^*$. In the usual way we may suppose $B \cap C = A$. Let $L$ be the underlying language. The three theories $T^*[A], T^*[B], T^*[C]$ are complete in the respective enrichments of $L$. Let $a$ be the enumeration of $A$, so that $L(A) = L(a)$. Then

$$T^*[A] \subseteq T^*[B] \cap L(a)$$

and this inclusion is an equality since $T^*[A]$ is complete. There is a similar observation for $C$. Thus

$$(B, a) \equiv (C, a)$$

which, by a standard argument, leads to an amalgamation

(Notice that this argument does not show $A \prec B$, unless $A$ is a model of $T^*$.)

(b) Given the conditions, it suffices to show that $T^*[A]$ is complete for each $A \models T$. For such a structure $A$ consider any pair of models $B, C$ of this enriched theory $T^*[A]$. Thus we have a wedge of embeddings, as above, where $B, C \in \mathcal{M}(T^*)$ but now $A \in \mathcal{M}(T)$. Since $T$ has $AP$, we may close this wedge

first to a model of $T$ and then extend this to a model $D$ of $T^*$. But $T^*$ is model complete, so both the upper embeddings are elementary, which leads to the required result.

(c) For each of the theories $T$ of Sections 5.2, 5.2, 5.2 the constructed theory $T^*$ is its model completion.
Of course, each model complete theory is its own model completion.

6.2 Consider elements $a, b$ of a commutative ring such that

$$ab \neq 0 \quad (ab)^2 = ab$$

hold. A simple induction gives

$$a^{n+1}b^{n+1} = ab$$

for each $n \in \mathbb{N}$, and hence $a^n \neq 0$.

This is enough to prove both Lemma 6.7 and Lemma 6.8((ii)⇒(i)).

F.2 Companion operators

6.3 By Definition 6.11(i) the two theories $T, T^b$ are companions, and hence $T^b = T^{ab}$ by 6.11(ii).

By 6.11(iii) we have

$$T^a \cap \forall_2 \subseteq T^{ab} = T^b$$

and hence $T^a \cap \forall_2 \subseteq T^b \cap \forall_2$. The converse holds by symmetry.

6.4 Consider and pair $T, T^*$ of companions theories with $T^*$ model complete. For each companion operator $(\cdot)^a$ we have

$$T^* \subseteq T^{*a} = T^a$$

since $T^*$ is $\forall_2$-axiomatizable and the two theories are companions. This inclusion shows that $T^a$ is also model complete. Thus $T^*$ and $T^a = T^{*a}$ are model complete companions, and hence $T^a = T^*$ by Corollary 6.3.

6.5 For each theory $T$ the companion $T^0$ is $\forall_2$-axiomatizable, and hence axiomatized by $T^0 \cap \forall_2$. For each companion operator $(\cdot)^a$ we have

$$T^0 \cap \forall_2 = T^a \cap \forall_2$$

by Exercise 6.3, and hence $T^0$ is axiomatized by $T^a \cap \forall_2$.

F.3 Existentially closed structures

6.6 Consider any pair

$$\mathcal{A} \subseteq \mathcal{B}$$

where $\mathcal{A} \in \mathcal{E}(T)$ and $\mathcal{B} \in \mathcal{S}(T)$. We have $\mathcal{B} \subseteq \mathcal{C}$ for some $\mathcal{C} \models T$ to give

$$\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{C}$$

and hence $\mathcal{A} \preceq_1 \mathcal{B}$ by the standard argument.

6.7 Suppose $\mathcal{E}(T)$ is elementary and let $T^* = \text{Th}(\mathcal{E}(T))$, so that $\mathcal{E}(T) = \mathcal{M}d(T^*)$. Since

$$\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} \preceq_1 \mathcal{B}$$
for \( A, B \in Md(T^*) \), we see that \( T^* \) is model complete. Also \( Md(T^*) = E(T) \subseteq S(T) \), so it suffices to show \( Md(T) \subseteq S(T^*) \). But this is an immediate consequence of the existence Theorem 6.21. 

6.8 Let \( \mathcal{U} \) be the union of a directed family \( A \subseteq E(T) \), consider \( \mathcal{U} \subseteq B \models T \), and suppose \( \mathcal{U} \models \phi(a) \) for some \( \forall_1 \)-formula \( \phi(v) \) and point \( a \) from \( \mathcal{U} \). There is some \( A \in A \) such that \( a \) comes from \( A \). But then \( \mathcal{U} \models \phi(a) \) and \( A \prec_1 B \), to give \( B \models \phi(a) \), as required. 

6.9 (a) Suppose 
\[
(A, a) \equiv (\exists_1) (B, b)
\]
so that we have a pair of embeddings 
\[
(A, a) \xrightarrow{f} (C, c) \quad (B, a) \xrightarrow{g} (C, c)
\]
where \( g \) is elementary. In particular, \( C \in S(T) \). But \( A \in E(T) \), so that \( f \) is a \( \prec_1 \)-embedding, and hence 
\[
(A, a) \equiv_1 (C, c) \equiv_1 (B, b)
\]
as required.

(b) Consider first \( A, B \in E(T) \) with \( A \subseteq B \). Then \( A \prec_1 B \) (since \( A \in E(T) \)), and hence 
\[
A \subseteq B \subseteq C \quad A < C
\]
for some structure \( C \). In particular, \( C \in S(T) \), so that \( B \prec_1 C \), and hence \( A \prec_2 B \).

Now suppose 
\[
A \equiv (\exists_1) B
\]
where \( A, B \in E(T) \). Up to isomorphism there is some \( C \in E(T) \) with \( A \subseteq C \) and \( B \subseteq C \). But now, by the observation above, we have \( A \prec_2 C \) and \( B \prec_2 C \), to give 
\[
A \equiv_2 C \equiv_2 B
\]
as required.

6.10 Let \( T^e = Th(E(T)) \). Since \( E(T) \subseteq S(T) \), we have \( T \cap \forall_1 \subseteq T^e \cap \forall_1 \). The converse follows by Theorem 6.21.

Lemma 6.19 shows that \( E(T) = E(T \cap \forall_1) \), and hence \( T^e \) depends only on \( T \cap \forall_1 \).

Finally, consider any \( A \in E(T) \). We have \( A \subseteq B \) for some \( B \models T \). But then \( A \prec_1 B \), and hence \( B \equiv (\forall_2) A \), so that \( A \models T \cap \forall_2 \). This shows that \( T \cap \forall_2 \subseteq T^e \). 

6.11 Consider any graph 
\[
A = (A, \sim)
\]
with \( A \in E(T) \) (where \( T \) is the theory of graphs). For finite disjoints subsets \( L, R \) of \( A \) let \( \theta(v) \) be the \( QF \)-formula 
\[
L \models v \models R
\]
using the abbreviated notation of Section 5.3. This formula contains parameters from \( A \) (namely, the nodes in \( L \cup R \)). By adjoining one extra node we obtain a graph \( B \) with 
\[
A \subseteq B \models (\exists v) \theta(v)
\]
and hence \( \models (\exists v)\theta(v) \) since \( \mathfrak{A} \prec_1 \mathfrak{B} \).

6.12 By Lemma 6.7 it suffices to show

\[
\mathcal{E}(T) \models (\forall v)[\bigwedge NNil(v) \rightarrow \text{nnil}(v)]
\]

holds. To this end consider any \( \mathfrak{A} \in \mathcal{E}(T) \) and any element \( a \) of \( \mathfrak{A} \) with \( \mathfrak{A} \models \text{Nnil}(a) \). This element is not nilpotent, so that Lemma 6.8 gives some \( \mathfrak{A} \subseteq \mathfrak{B} \models T \) with \( \mathfrak{B} \models \text{nnil}(a) \). But now \( \mathfrak{A} \prec_1 \mathfrak{B} \) and hence \( \mathfrak{A} \models \text{nnil}(a) \), as required.

This does not show

\[
T^e \cup \text{NNil}(v) \vdash \text{nnil}(v)
\]

for the argument above works only for \( \mathfrak{A} \in \mathcal{E}(T) \), not for arbitrary \( \mathfrak{A} \models T^e \).

F.4 Existence of existentially closed structures

6.13 This is proved by a modification of the argument used in the second part of the proof of Theorem 6.23. To do that we need a slight refinement.

Let \( \mathcal{G}(T), \mathcal{G}'(T) \) be two classes with the indicated properties. We first show that

\[
[n] \quad \text{We have}
\]

\[
\mathfrak{A} \subseteq \mathfrak{A}' \implies \mathfrak{A} \prec_{2n} \mathfrak{A}'
\]

for each \( \mathfrak{A} \in \mathcal{G}(T), \mathfrak{A}' \in \mathcal{G}'(T) \) by induction on \( n \).

The base case \([0]\) is trivial.

For the induction step \( n \mapsto n + 1 \), suppose

\[
\mathfrak{A} \subseteq \mathfrak{A}'
\]

for some \( \mathfrak{A} \in \mathcal{G}(T) \) and \( \mathfrak{A}' \in \mathcal{G}'(T) \). By properties \((i)\) and \((i')\) we have

\[
\mathfrak{A} \subseteq \mathfrak{A}' \subseteq \mathfrak{B} \subseteq \mathfrak{B}'
\]

for some \( \mathfrak{B} \in \mathcal{G}(T) \) and \( \mathfrak{B}' \in \mathcal{G}'(T) \). The properties \((ii)\) and \((ii')\) give

\[
\mathfrak{A} \prec \mathfrak{B}' \quad \mathfrak{A}' \prec \mathfrak{B}'
\]

and we have

\[
\mathfrak{B} \prec_{2n} \mathfrak{B}'
\]

by the induction hypothesis \([n]\). This gives first

\[
\mathfrak{A}' \prec_{2n+1} \mathfrak{B}
\]

and then

\[
\mathfrak{A} \prec_{2n+2} \mathfrak{A}'
\]

for the required result.
Now consider any $\mathfrak{A} \in G(T)$. By $(i')$ we have
\[ \mathfrak{A} \subseteq \mathfrak{A}' \]
for some $\mathfrak{A}' \in G'(T)$. But now the preliminary observation gives
\[ \mathfrak{A} \prec \mathfrak{A}' \]
and hence $\mathfrak{A} \in G'(T)$ by $(iii')$.
This shows that $G(T) \subseteq G'(T)$, and $G'(T) \subseteq G(T)$ holds by symmetry.

6.14 (a) Consider any $\mathfrak{A} \in G(T)$. By the properties of $E(T)$ and $G(T)$ we have
\[ \mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{C} \prec - \]
for some $\mathfrak{B} \in E(T)$ and some $\mathfrak{C} \in G(T)$. But now we have $\mathfrak{A} \prec \mathfrak{B}$, and hence $\mathfrak{A} \in E(T)$, as required.
(b) Let $\mathcal{A}$ be a directed subfamily of $G(T)$ with union $\mathfrak{B}$. We have $\mathfrak{B} \in S(T)$, and hence $\mathfrak{B} \subseteq \mathfrak{C}$ for some $\mathfrak{C} \in G(T)$. For each $\mathfrak{A} \in \mathcal{A}$ we have
\[ \mathfrak{A} \prec \mathfrak{B} \prec \mathfrak{C} \]
the first by the standard properties of directed systems and the second by the properties of $G(T)$. These give $\mathfrak{B} \prec \mathfrak{C}$, and hence $\mathfrak{B} \in G(T)$, as required.
(c) We modify the argument of Solution 6.9
We are given
\[ (\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b) \]
where $\mathfrak{A}, \mathfrak{B} \in G(T)$. In the standard way we produce a pair of embeddings
\[ (\mathfrak{A}, a) \xrightarrow{f} (\mathfrak{C}, c) \quad (\mathfrak{B}, a) \xrightarrow{g} (\mathfrak{C}, c) \]
where, in the first instance, $\mathfrak{C} \in S(T)$. By extension we may assume $\mathfrak{C} \in G(T)$, and hence both $f$ and $g$ are elementary embeddings. This gives
\[ (\mathfrak{A}, a) \equiv (\mathfrak{C}, c) \equiv (\mathfrak{B}, b) \]
for the required result.

6.15 (a) We modify Solution 6.10.
Let $T^g = Th(G(T))$. Since $G(T) \subseteq S(T)$, we have $T \cap \forall_1 \subseteq T^e \cap \forall_1$. The converse follows by property $(i)$ of $G(T)$.
Let $T_1, T_2$ be a pair of companion theories. Since $S(T_1) = S(T_2)$ we see that each of $G(T_1), G(T_2)$ satisfies the conditions $(i, ii, iii)$ for both $T_1$ and $T_2$. Thus the uniqueness property of Exercise 6.13 shows that $G(T_1) = G(T_2)$, and hence $T^g_1 = T^g_2$.
Finally, since $G(T) \subseteq \mathcal{E}(T)$, we have $T \cap \forall_2 \subseteq T^e \subseteq T^g$, as required.
(b) Suppose first that $T$ has a model companion $T^*$. Then $Md(T^*)$ has the three characterizing properties $(i, ii, iii)$ of $G(T)$, and hence is this class.
The converse follows by the argument of Solution 6.6.
F.5 The use of types

6.16 As indicated we consider some $\mathfrak{A} \in S(T)$.

(i) $\Rightarrow$ (ii). Suppose $\mathfrak{A} \in E(T)$, consider any point $a$ of $\mathfrak{A}$ and let $\Sigma(v)$ be the $\exists_1$-type of $a$ in $\mathfrak{A}$. Consider any $\exists_1$-formula $\theta(v)$ such that

$$T \cup \Sigma(v) \cup \{\theta(v)\}$$

is consistent. We must show that $\theta \in \Sigma$.

Since the set of formulas is consistent we have

$$\mathfrak{B} \models T \quad \mathfrak{B} \models \Sigma(b) \quad \mathfrak{B} \models \theta(b)$$

for some structure $\mathfrak{B}$ and point $b$. The second of these give

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b)$$

and hence

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

since $\mathfrak{A} \in E(T)$. Since $\mathfrak{B} \models \theta(b)$, this gives $\mathfrak{A} \models \theta(a)$, and hence $\theta \in \Sigma$, as required.

(ii) $\Rightarrow$ (i). Suppose the structure $\mathfrak{A}$ has the indicated property. Consider any structure $\mathfrak{A} \subseteq \mathfrak{B} \models T$ and suppose $\mathfrak{B} \models \theta(a)$ for some $\exists_1$-formula $\theta(v)$ and point $a$ of $\mathfrak{A}$. We must show that

$$\mathfrak{A} \models \theta(a)$$

in other words that $\theta(v)$ belongs to the $\exists_1$-type $\Sigma(v)$ of $a$ in $\mathfrak{A}$.

The structure $\mathfrak{B}$ ensures that

$$T \cup \Sigma(v) \cup \{\theta(v)\}$$

is consistent, and hence $\theta \in \Sigma$ by the given maximality of $\Sigma$. \hfill \blacksquare

6.17 Suppose first that $\mathfrak{A} \in G(T)$ and that $\mathfrak{A} \models \phi(a)$ for some formula $\phi(v)$ and point $a$ of $\mathfrak{A}$. Let $\Sigma(v)$ be the $\exists_1$-type of $a$ in $\mathfrak{A}$. Since $\mathfrak{A} \models \Sigma(a)$ it suffices to show that

$$G(T) \models (\forall v)[\bigwedge \Sigma(v) \rightarrow \phi(v)]$$

holds.

Consider any $\mathfrak{B} \in G(T)$ and any point $b$ of $\mathfrak{B}$ with $\mathfrak{B} \models \Sigma(b)$. We have

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b)$$

and hence

$$(\mathfrak{A}, a) \equiv (\mathfrak{B}, b)$$

by Solution 6.14(c). Since $\mathfrak{A} \models \phi(a)$, this gives $\mathfrak{B} \models \phi(b)$, for the required result.
F.5. The use of types

Conversely, suppose $\mathfrak{A} \in \mathcal{S}(T)$ has the indicated property. Consider any $\mathfrak{B} \in \mathcal{G}(T)$ with $\mathfrak{A} \subseteq \mathfrak{B}$. The assumed property of $\mathfrak{A}$ gives $\mathfrak{A} \prec \mathfrak{B}$, and hence $\mathfrak{A} \in \mathcal{G}(T)$. ■

6.18 Suppose first that $\mathfrak{A} \in \mathcal{E}(T)$ and consider any $\forall_1$-formula $\phi(v)$ and any point $a$ of $\mathfrak{A}$ which matches the batch $v$. We must show that $\mathfrak{A} \models \Omega(T, \phi)(a)$ does not hold.

If $\mathfrak{A} \models \neg \phi(a)$ then we are done. Thus we may suppose $\mathfrak{A} \models \phi(a)$. By Theorem 6.26 this gives $\mathfrak{A} \models \theta(a)$ for some $\forall_1$-formula $\theta(v)$. Since $\neg \theta \in \Omega(T, \phi)$, we are done.

Conversely, suppose $\mathfrak{A} \in \mathcal{S}(T)$ omits $\Omega(T, \phi)$ for each $\forall_1$-formula $\phi$. We use Theorem 6.26 to show that $\mathfrak{A} \in \mathcal{E}(T)$.

To his end suppose $\mathfrak{A} \models \phi(a)$ for some $\forall_1$-formula $\phi(v)$ and point $a$ of $\mathfrak{A}$. Since $\mathfrak{A}$ omits $\Omega(T, \phi)$ and $\mathfrak{A} \models \phi(a)$ we must have $\mathfrak{A} \models \theta(a)$ for some $\exists_1$-formula $\theta(v)$ where $T \vdash \theta \rightarrow \phi$ holds. This verifies condition (iii) of Theorem 6.26.

In truth the condition ‘$\mathfrak{A}$ omits $\Omega(T, \phi)$’ is little more than a rephrasing of condition (iii) of Theorem 6.26. ■
Large and small structures

7.1 Given $A \in E(T)$ there is some $M \in U(T)$ with $A \subseteq M$ (since $U(T)$ is cofinal in $S(T)$). But $A \in E(T)$ and hence $A \prec_1 M$.

Conversely, suppose $A \prec_1 M$ for some $M \in U(T)$. Since $U(T) \subseteq E(T)$ this gives $A \in E(Y)$ by one of the characteristic properties of $E(T)$.

7.2 For the given $A \in S(T)$ consider any $M \in U(T)$ with $A \subseteq M$. We show

$$[n] \ A \prec_{n+1} M$$

for each $n \in \mathbb{N}$. We do not need a proof by induction.

Fix $n \in \mathbb{N}$. We have a wedge, as to the left, where $f$ is the embedding under consideration and $g$ is the given $\prec_{n+1}$-embedding.

This wedge can be completed to a commuting square, as on the right, where in the first instance $h$ is a $\prec$-embedding. In particular, that $\mathcal{L} \in S(T)$, and hence we may embed this into a new $\mathcal{L} \in U(T)$. With this new $\mathcal{L}$ both $h, k$ are $\prec$-embeddings, and hence $f$ is a $\prec_{n+1}$-embedding, as required.

7.3 (a) This is just Exercise 7.1.

(b) This is trivial.

(c) By the solution to Exercise 7.5 we have

$$A \subseteq M \implies A \prec_{n+1} M$$

for each $A \in E_n(T)$ and $M \in U(T)$. This gives $(i) \implies (ii)$. The implications $(ii) \implies (iii)$ and $(iii) \implies (i)$ are trivial.

7.4 We use the characterization of $G(T)$ given by Theorem 7.13.

Suppose first that $T$ has a model companion $T^*$. Then $Md(T^*)$ satisfies the three characteristic properties (i, ii, iii), and hence $G(T) = Md(T^*)$ is elementary.
Conversely, suppose $G(T)$ is elementary and let $T^* = Th(G(T))$. We have $G(T) = Md(T^*)$. Property (i) shows that $T$ and $T^*$ are companions, and property (i) shows that $T^*$ is model complete. Thus $T^*$ is the model companion of $T$. ■

7.5 Let $A$ be a directed system taken from $G(T)$ and let $\mathfrak{A} = \bigcup A$. Using Theorem 7.13 we have $\mathfrak{A} \simeq \mathfrak{A}$ for each $\mathfrak{A} \in A$. To show that $\mathfrak{A} \in G(T)$ we use the characterization given by Theorem 7.15.

Suppose $\mathfrak{A} \models \phi(a)$ for some formula $\phi(v)$ and some point $a$ of $\mathfrak{A}$. This point lives in some $\mathfrak{A} \in A$, and then $\mathfrak{A} \models \phi(a)$ since $\mathfrak{A} \simeq \mathfrak{A}$. Since $\mathfrak{A} \in G(T)$ we have

$$\mathfrak{A} \models \Theta(a) \quad G(T) \models (\forall v)[\Theta \rightarrow \phi]$$

for some $\exists_1$-type $\Theta(v)$. The left hand condition gives $\mathfrak{A} \models \Theta(a)$, and hence $\mathfrak{A} \in G(T)$. ■

G.3 For §7.3 - to be done

G.4 Atomicity

7.6 Suppose first that $\theta$ is complete over $T$ and that both

$$T \cup \{\theta, \psi_1\} \quad T \cup \{\theta, \psi_2\}$$

are consistent for formulas $\psi_1, \psi_2$. Then neither of

$$T \vdash \theta \rightarrow \neg \psi_1 \quad T \vdash \theta \rightarrow \neg \psi_2$$

hold, and hence both

$$T \vdash \theta \rightarrow \psi_1 \quad T \vdash \theta \rightarrow \psi_2$$

hold (since $\theta$ is complete over $T$). But $\theta$ is consistent over $T$, and hence $T \cup \{\psi_1, \psi_2\}$ is consistent.

Conversely, suppose $\theta$ has the indicated property. For an arbitrary formula $\psi$ we can have at most one of

$$T \vdash \theta \rightarrow \psi \quad T \vdash \theta \rightarrow \neg \psi$$

since $\theta$ is consistent with $T$. By way of contradiction, suppose that neither of these hold. Then both

$$T \cup \{\theta, \neg \psi\} \quad T \cup \{\theta, \psi\}$$

are consistent, so that

$$T \cup \{\neg \psi, \psi\}$$

is consistent, which is nonsense. ■

7.7 Observe that the formal Definition 7.23 and the implication of the exercise are contrapositives. ■

7.8 Consider any $\forall_1$-formula $\phi$ which is consistent with $T^0$. Since $T^0$ is 0-complete, there is a $\exists_1$-formula $\psi$ which is consistent with $T^0$ and with $T^0 \vdash \psi \rightarrow \phi$. This formula $\psi$ is consistent with $T$, which is given to be $\exists_1$-atomic, so there is some formula $\theta$ which is
∃₁-complete over $T$ and with $T^0 \vdash \theta \rightarrow \psi$. Thus we have $T^0 \theta \rightarrow \phi$ and hence $T\theta \rightarrow \phi$, as required. 

7.9 (a) An $∃_{n+1}$-formula $\theta$ is $∃_{n+1}$-complete over a theory $T$ if it is consistent with $T$ and

$$T \cup \{ \theta, \psi_1 \} \text{ is consistent} \implies T \cup \{ \psi_1, \psi_2 \} \text{ is consistent}$$

holds for all $∃_{n+1}$-formulas $\psi_1, \psi_2$ (with $\partial \psi_1 \cup \partial \psi_2 \subseteq \partial \theta$). Equivalently, using Exercise 7.7, an $∃_{n+1}$-formula $\theta$ is $∃_{1}$-complete over $T$ if

$$T \vdash \phi_1 \lor \phi_2 \implies T \vdash \theta \rightarrow \phi_1 \text{ or } T \vdash \theta \rightarrow \phi_2$$

holds for all $∀_{n+1}$-formulas $\phi_1, \phi_2$.

(b) Consider a formula $\theta$ which is $∃_{n+1}$-complete over $T$ for all large $n$. In particular, $\theta$ is consistent with $T$. Consider any formula $\psi$ (with $\partial \psi \subseteq \partial \theta$). Since $\theta$, is consistent with $T$, at most one of

$$T \vdash \theta \rightarrow \psi \quad T \vdash \theta \rightarrow \lnot \psi$$

holds, so it suffices to show that at least one holds. We may chose $n$ large enough so that $\theta$ is a $∃_{n+1}$-formula, and both $\psi, \lnot \psi$ are $∀_{n+1}$-formulas. Since

$$T \vdash \psi \lor \lnot \psi$$

the required result follows by the variant of the official definition.

Conversely, suppose that $\theta$ is a $∃_{n+1}$-formula which is complete over $T$. Suppose

$$T \vdash \phi_1 \lor \phi_2$$

where both $\phi_1$ and $\phi_2$ is a $∀_{n+1}$-formula. Since $\theta$ is complete over $T$ at least one of

$$T \vdash \theta \rightarrow \phi_1 \quad T \vdash \theta \rightarrow \lnot \phi_1$$

holds. If then second one holds, then so does

$$T \vdash \theta \rightarrow \phi_2$$

(since $T \vdash \phi_1 \lor \phi_2$) to show that $\theta$ is $∃_{n+1}$-complete over $T$. 

7.10 To be done if necessary 

7.11 This is needed to ensure that the required b&f system is non-empty. 

7.12 This is a refinement of the proof of Theorem 7.27. Let $\mathfrak{A}$ be the unique countable $∃_1$-atomic structure for $T$. Consider and model $\mathfrak{B} \models T^0$.

For the time being suppose $[Check this is right]$

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b)$$

for points $a$ from $\mathfrak{A}$ and $b$ from $\mathfrak{B}$. Since $\mathfrak{A} \in E(T)$, this ensures that

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$
holds. Consider any element $x$ of $\mathfrak{A}$. As in the proof of Theorem 7.27, there is some element $y$ of $\mathfrak{B}$ such that

$$(\mathfrak{A}, a \models x) \equiv (\exists_1) (\mathfrak{B}, b \models y)$$

holds.

Now let $a$ be a full enumeration of $\mathfrak{A}$. By iterated use of the above observation we produce a partial enumeration $b$ of $\mathfrak{B}$ such that

$$(\mathfrak{A}, a) \equiv (\exists_1) (\mathfrak{B}, b)$$

holds, and hence

$$(\mathfrak{A}, a) \equiv_1 (\mathfrak{B}, b)$$

(since $\mathfrak{A} \in \mathcal{E}(T)$). This gives the required embedding. ■

7.13 Left $\mathfrak{A}$ be the unique countable structure which is $\exists_1$-atomic for $T$. Since $\mathcal{A}(T) \subseteq \mathcal{E}(T)$, we have $T^* = T^e \subseteq T^a$. Consider any $\mathfrak{B} \models T^*$. By Exercise 7.11 we have $\mathfrak{A} \subseteq \mathfrak{B}$, and hence $\mathfrak{A} \prec \mathfrak{B}$ (since $T^*$ is model complete), and hence $\mathfrak{B} \models T^a$, to show that $T^* = T^a$. ■